



BUDAPESTI MŰSZAKI ÉS GAZDASÁGTUDOMÁNYI EGYETEM

Gazdaság- és Társadalomtudományi Kar

Testnevelési Központ

A dinamikus pontrendszerek szerepe és alkalmazása a labdarúgó-bajnokságokban

Készítette: Bologa Eduárd
Mérnökinformatikus szakos
hallgató

Konzulens: Bartha Zsolt
Testnevelő Tanár,
igazgatóhelyettes BME GTK,
Testnevelési Központ
Rácz László
Tudományos munkatárs

**Országos Tudományos Diákköri Konferencia
Budapest, 2025**

Tartalomjegyzék

1. Bevezetés.....	1
1.1. A dinamikus pontrendszerek szerepe a sportban	1
1.2. Problémafelvetés.....	1
2. Irodalmi áttekintés.....	2
2.1. Az Élő pontrendszer áttekintése.....	2
2.2. A döntetlenek kezelése a pontszámítási rendszerekben.....	4
2.3. A dinamikus pontszámítás fejlődése és szerepe	5
3. Célkitűzés és hipotézis felállítása.....	7
4. Módszer	8
4.1. Kutatási környezet és modellezési eszközök	8
4.2. Használt adatkészlet és kiválasztott modellek	8
4.2.1. Baseline Élő modell	8
4.2.2. Uniform dinamikus Élő modell	8
4.2.3. Végső dinamikus Élő modell	8
4.3. Paraméterek és egyedi beállítások	9
4.4. Statisztikai elemzés	9
5. Eredmény.....	11
5.1. Rangsorolási modellek eredményei	11
5.2. Statisztikai korrelációk és Tau-metrikák értelmezése.....	11
5.3. Győzelmi arányok az Élő pontkülönbség alapján.....	14
5.4. Becslés a mérkőzések kimenetelére	14
6. Kitekintés és konklúzió	15
6.1. A modell továbbfejlesztése és finomítása.....	15
6.2. Felhasználási lehetőség.....	15
6.3. Konklúzió.....	15
7. Irodalomjegyzék	16

ABSZTRAKT

A labdarúgó-bajnokságok hagyományos pontszámítási rendszere régóta szolgálja a csapatok rangsorolását, azonban a statisztikai és adatfeldolgozási módszerek fejlődésével, valamint a sportinnováció előrehaladásával egyre nagyobb igény mutatkozik a dinamikus pontrendszerek bevezetésére. Az ilyen rendszerek lehetővé teszik a csapatok teljesítményének folyamatos, mérkőzésről-mérkőzésre történő kiigazítását, ami pontosabb képet adhat a csapatok erőviszonyairól a szezon folyamán.

Dolgozatom célja egy olyan módosított Élő-féle rangsorolási modell kifejlesztése és bemutatása, amely valós időben értékeli a csapatok teljesítményét a szezon során. Az Élő modell eredetileg kétkimenetelű játékokhoz (győzelem-vereség) készült, és nem kezelte megfelelően a döntetlent, amely kulcsfontosságú a háromkimenetelű játékokban, mint a labdarúgás. Ezt az esetet részben kezeli az Élő-Davidson modell, így ennek megfelelően a kutatásom egy újabb modell megalkotását tervezi. Ezen megközelítés egyik legfontosabb eleme a döntetlenek "értékének" még pontosabb leképezése, amely biztosítja a rangsorolás megbízhatóságát és pontosságát.

A rendszer figyelembe veszi a mérkőzések különböző szempontjait, beleértve az aktuális formát, az ellenfelek relatív erősségét, a pályaválasztást, valamint a korábbi teljesítményt. Az ilyen dinamikus rangsorolási rendszerek alkalmazásával egy pontosabb, reálisabb versenykép alakulhat ki, amely nemcsak a bajnoki végeredmény megállapításában, hanem a csapatok erőforrás-tervezésében és stratégiai döntéseiben is hasznos lehet.

A dolgozatban bemutatott modellek és számítási módszerek tesztelése különböző szimulációkon keresztül történik, amelyek célja a pontosság és a rendszer megbízhatóságának elemzése. A kísérleti eredmények alapján javaslatokat fogalmazok meg a dinamikus pontrendszerek széleskörű alkalmazására, különös tekintettel a labdarúgó-bajnokságok speciális igényeire.

Kulcsszavak:

sportinnováció, dinamikus pontrendszer, labdarúgó-bajnokság, Élő algoritmus, rangsorolás, adatfeldolgozás, statisztikai rangkorreláció

1. BEVEZETÉS

1.1. A dinamikus pontrendszerek szerepe a sportban

A versenysportok területén az Élő-rangsorolási rendszer hatékonyan alkalmazható csapatok, vagy akár játékosok rangsorolására, amint azt a labdarúgásban és más sportágakban végzett kutatások is bizonyították. Az Élő algoritmus például erős eszköznek bizonyult a mérkőzések kimeneteleinek előrejelzésére, és számos esetben felülmúlta a hagyományos rangsorolási módszereket [1][2]. Ez a szemléletes előrejelző képesség annak köszönhető, hogy a rendszer képes figyelembe venni számos olyan tényezőt, mint például a visszamenőleges teljesítményadatok, a csapat aktuális formája és a két csapat közötti erőviszonyok relatív különbsége, amelyek befolyásolhatják a teljesítményt.

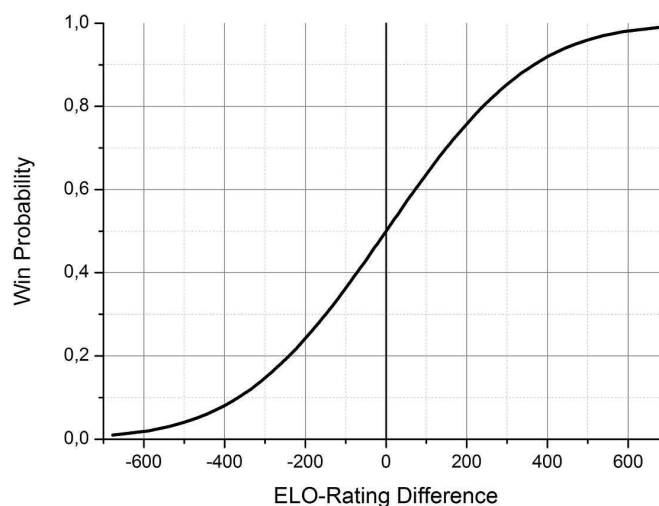
1.2. Problémafelvetés

A labdarúgó-bajnokságokban leginkább elterjedt hagyományos pontszámítási rendszerek, amelyek fix 0, 1 vagy 3 pontot osztanak ki a mérkőzések eredményei alapján, hosszú ideje a labdarúgó bajnokságok pontszámítási alapját képezik. Ezek a rendszerek azonban nem képesek teljes mértékben visszatükrözni a csapatok teljesítményének valós dinamikáját. A hagyományos rangsorolási rendszerek gyakran figyelmen kívül hagyják a mérkőzések számos kritikus tényezőjét, mint például a csapatok aktuális formája, az ellenfelek erőviszonyai, a hazai pálya előnye, valamint a sérülések vagy egyéb, a csapat teljesítményét befolyásoló körülmények. Ez különösen fontos a labdarúgásban, ahol a döntetlen is a rangsorolási rendszer szerves részét képezi, szemben más sportágakkal, mint például a kosárlabda vagy a tenisz [3].

2. IRODALMI ÁTTEKINTÉS

2.1. Az Élő pontrendszer áttekintése

Az **Élő-rangsorolási rendszert** eredetileg Élő Árpád, magyar származású amerikai fizikus és profi sakkozó, fejlesztett ki 1978-ban, a sakkjátékosok rangsorolási céljából. Széles körben használják, a mai napig, különböző területeken a versenyeredmények alapján történő teljesítményértékelésre és rangsorolásra. Alapelve, hogy a résztvevők értékelését a mérkőzéseik eredményei alapján frissíti, tükrözve ezzel a játékosok relatív tudásszintjét. Minden résztvevő egy alapértékeléssel indul, amely jellemzően 1000 pont, és ez az értékelés minden mérkőzés után az eredmény függvényében módosul. A módosítás mértéke leginkább attól függ, hogy az előző értékelések alapján mekkora volt a győzelem valószínűsége [4][5].



1. ábra: A győzelem valószínűsége az Élő rendszerben az értékelési különbség függvénye

Forrás: <https://medium.com/@cricketscience/how-do-sports-ratings-work-9433f1e6c015>

Az Élő rendszer folyamatos jellege lehetővé teszi olyan értékelések létrehozását, amelyek tükrözik az alanyok relatív tudásszintjét, így különösen hasznos olyan helyzetekben, ahol a teljesítmény változékony, és számos tényező befolyásolja. Ezt a tulajdonságot alátámasztják azok a tanulmányok, amelyek kiemelik az interakciók gyakoriságának és a hierarchiák meredekségének szerepét a dominancia rangsorok pontos meghatározásában [6]. Emellett a rendszer páros összehasonlításokon alapul, ami csökkenti az elfogultságot, és növeli a rangsorok megbízhatóságát [7].

Az Élő pontrendszer matematikai háttere:

Jelölje N a játékosok számát ($N = 1, 2, \dots, n$).

Az egyes mérkőzések kimenetelét y_t jelöli, ahol a mérkőzést 2 csapat játssza t pillanatban.

Jelölje i_t a "hazai" csapatot és j_t a „vendég” csapatot.

Ezek után meghatározzuk a 3 lehetséges mérkőzés kimenetelt:

- Győzelem a hazai csapat számára:

$$y_t = H, \text{ ha } i_t > j_t \quad (1)$$

- Döntetlen:

$$y_t = D, \text{ ha } i_t = j_t \quad (2)$$

- Győzelem a vendég csapat számára:

$$y_t = A, \text{ ha } i_t < j_t \quad (3)$$

A kényelmesebb jelölés érdekében ezeket a kategóriákat bináris változókra bontom fel, majd egy indikátorfüggvény segítségével kifejezem őket:

- h_t – azt jelzi, hogy a hazai csapat nyert

$$h_t = I[y_t = H] \quad (4)$$

- d_t – azt jelzi, hogy az eredmény döntetlen lett

$$d_t = I[y_t = D] \quad (5)$$

- a_t – azt jelzi, hogy a vendégcsapat nyert

$$a_t = I[y_t = A] \quad (6)$$

Az indikátorfüggvényekre igaz, hogy $I[H] = 1$, ha H igaz, és $I[H] = 0$, ha H hamis. Hasonlóképpen igaz az A és a D kifejezésre is az egyenlet.

Ez a három kimenetel kölcsönösen kizárja egymást, tehát fennáll a következő egyenlet:

$$h_t + d_t + a_t = 1 \quad (7)$$

Jelölje Y_t valószínűségi változó a mérkőzések eredményeit, a lehetséges kimenetek halmaza pedig $(\{H, D, A\})$.

$y_1, y_2, \dots, y_t$ mérkőzések eredményeit megfigyelve, minden csapathoz egy numerikus értéket rendelünk, amit $\Theta_{t,n}$ jelöl és megfelel az adott csapat értékelési szintjének.

A valószínűségi modell (melynek szerepe, hogy egy jelenséget vagy folyamatot valószínűségek segítségével írjon le), amely összekapcsolja a mérkőzés eredményét a résztvevő csapatok értékelési szintjeivel a következő lineáris modellel adható meg:

$$P\{i > j \mid \Theta_i, \Theta_j\} = F_H(\Theta_i - \Theta_j) \quad (8)$$

$$P\{i < j \mid \Theta_i, \Theta_j\} = F_A(\Theta_i - \Theta_j) \quad (9)$$

$$F_H(z) = F(z) \quad (10)$$

$$F_A(z) = F(-z) = 1 - F(z) \quad (11)$$

Ahol $F_H(z)$ egy monoton növekvő, $F_A(z)$ egy monoton csökkenő függvény; z megegyezik két csapat értékelési különbségével:

$$z = \Theta_i - \Theta_j \quad (12)$$

A modell szimmetrikus és a teljes valószínűség törvényét követi.

A logisztikus függvény, amelyet az Élő pontok meghatározására használnak:

$$F(z) = \frac{1}{1 + 10^{-\frac{z}{\sigma}}} = \frac{10^{\frac{0.5z}{\sigma}}}{10^{\frac{0.5z}{\sigma}} + 10^{-\frac{z}{\sigma}}} \quad (13)$$

Ahol $\sigma > 0$ egy skála paraméter, amely meghatározza, hogy az értékelési különbség milyen mértékben befolyásolja a győzelem valószínűségét. [8]

2.2. A döntetlenek kezelése a pontszámítási rendszerekben

A döntetlenek kérdése a versenysportokban, különösen a labdarúgó-bajnokságokban, komoly kihívást jelent a rangsorolási rendszerek számára [8]. A hagyományos pontozási rendszerek, mint például a győzelemért járó három pont és a döntetlenért járó egy pont, ellentmondásokat eredményezhetnek a csapatok teljesítményének és versenyképességének értékelésében. Ez végképpen szembetűnő a hazai pálya előnye kontextusában, ahol az értékelési módszerek gyakran nem veszik megfelelően figyelembe a döntetlenek hatását az összesített rangsorokra. García és munkatársai például rámutatnak, hogy a hazai pálya előnyének értékelése torzulhat, ha a döntetlenek nincsenek megfelelően beépítve az elemzésbe, ami félrevezető következtetésekhez vezethet a csapatok teljesítményével kapcsolatban [9].

A labdarúgó-bajnokságok dinamikus jellege, ahol a csapatok teljesítménye mérkőzésről mérkőzésre változhat, tovább növeli a döntetlenek által felvetett problémák komplexitását. Ezért kiemelten fontos egy olyan rangsorolási rendszer, amely képes alkalmazkodni ezekhez a változásokhoz, és pontosan képes kezelni a döntetlenekből adódó anomáliákat. Számos kutatás bizonyította, hogy a hagyományos rangsorolási módszerek gyakran nem tudják megfelelően kezelni a csapatok teljesítményének bonyolultságát a döntetlenek jelenlétében, és egyre nagyobb igény mutatkozik a dinamikus pontozási rendszerek, valamint a fejlett statisztikai modellek alkalmazására különböző területeken [10].

2.3. A dinamikus pontszámítás fejlődése és szerepe

Az előző pontban említett kihívásnak a kezelésére a meglévő Élő-rangsorolási rendszert szükséges finomítani és továbbfejleszteni, mivel az eredeti rendszer a döntetlenek kezelésére egy statikus valószínűségi paramétert ($p = 0,5$) alkalmazott, amely nem feltétlenül tükrözi megfelelően a mérkőzések valós kimeneteleit. Ennek kiküszöbölése érdekében a dolgozat az **Élő-Davidson algoritmust** alkalmazza, amely bevezet egy új változót és pedig a κ (kappa) egyenlőségi paramétert, így pontosabb és megbízhatóbb adatokat nyújt a döntetlenek kezelésére.[11]

A döntetlenek modellezése a következőképpen történik:

$$F_H(z) = F(z; \kappa) = \frac{10^{\frac{0.5z}{6}}}{10^{\frac{0.5z}{6}} + 10^{\frac{z}{6} + \kappa}} \quad (14)$$

$$F_A(z) = F(-z; \kappa) = \frac{10^{\frac{0.5z}{6}}}{10^{\frac{0.5z}{6}} + 10^{\frac{z}{6} + \kappa}} \quad (15)$$

$$F_D(z) = \kappa \sqrt{F_H(z) - F_A(z)} = \frac{\kappa}{10^{\frac{0.5z}{6}} + 10^{\frac{z}{6} + \kappa}} \quad (16)$$

Ahol $\kappa > 0$ az egyenlőségi paraméter, amit meg kell határozni, és a csapatok pontszámainak különbsége alapján dinamikusan változik. Egy előre meghatározott szorzó hatására a gyengébb csapat κ értéke 0,1-re, míg az erősebb csapaté 0,025-re áll be.

A csapatok értékelési szintjének változásának meghatározására a következő képletet használom:

$$\Theta'_{t+1,i} = \Theta'_i + K[s_i - F(\Delta_i; \kappa)] \quad (17)$$

Ahol Δ_i az értékelési különbséget jelenti a két csapat között, K az adaptációs lépésköz (súlyozási tényező), amely meghatározza, hogy a pontszámok mennyire változnak egy mérkőzés után, $s_i \in \{0, 1/2, 1\}$ (0, ha a csapat veszít, 1, abban az esetben ha nyer a csapat és $1/2$ ha döntetlen születik).

A K súlyozási tényező meghatározása:

- 30 Bajnokok Ligája mérkőzések esetén
- 20 minden más bajnokság, torna esetében,
- 10 barátságos mérkőzésekre.

A gólkülönbségek szerepe a pontozási rendszerben a következőképpen befolyásolja az eredményeket:

$$K_{új} = K * (1 + 0.75 + \frac{M-3}{8}) \quad (18)$$

ahol:

- $K_{új}$ az új, módosított súlyozási tényező, amely a gólkülönbség figyelembevételével van frissítve
- K az eredeti súlyozási tényező
- M : A mérkőzés gólkülönbsége (pl. ha a csapat 4-0-ra nyert, akkor $M=4$)
- A 0.75-ös faktor empirikusan megválasztott érték, amely megfelelő mértékű különbséget jelent a szoros és a jelentős különbséggel elért győzelmek között.

A pontozási rendszer továbbá figyelembe fogja venni, hogy a mérkőzés hazai vagy idegenbeli pályán került megrendezésre, és a következő módon számol vele:

$$HFA = HFA + 0.0075 * \sum \Delta\Theta \quad (19)$$

Ahol a HFA -val jelölöm a hazai pálya előnyét, és a $\sum \Delta\Theta$ pedig a csapatok értékelésbeli különbségét adja meg.

Ezt is figyelembe véve így fejlesztettem tovább a modellt:

$$\Delta\Theta = (\Theta_A - \Theta_B) + HFA \quad (20)$$

$$F(z) = \frac{1}{1 + 10^{-\frac{\Delta\Theta}{6}}} \quad (21)$$

3. CÉLKITŰZÉS ÉS HIPOTÉZIS FELÁLLÍTÁSA

A kutatási hipotéziseim feltételezik, hogy az újonnan kidolgozott Élő-alapú rangsorolási modell pontosabban tükrözi a csapatok valós teljesítményét és aktuális erősségét a szezon végére, mint a hagyományos pontozási rendszer.

Hipotézis 1: Az általam megalkotott dinamikus modell a szezon során jobban tudja sejtetni a mérkőzések végkimenetelét, mivel meccsenként képes alkalmazkodni a csapatok közti erőviszonyok változásához és ezáltal nagyobb prediktív tulajdonsággal rendelkezik, és hasznosabb a csapatok relatív erősségének becslésében, mint a hagyományos rangsor.

Hipotézis 2: A dinamikus komponensekkel bővített Élő-alapú rangsorolási modell realisabb képet nyújt a csapatok szezon közbeni formaváltozásairól mint a klasszikus rangsorolási módszer, ideértve a hullámvölgyeket, győzelmi és vereségi sorozatukat. A modell pontosabban követi az idény során bekövetkező teljesítmény-ingadozásokat, és képes gyorsan alkalmazkodni a csapatok erőviszonyainak alakulásához.

4. MÓDSZER

A módszertani fejezet célja, hogy részletesen bemutassam, hogyan került kidolgozásra a rangsorolási modell, milyen adatokat használtam, és hogyan történt a klasszikus és az új rangsorok összehasonlítása.

4.1. Kutatási környezet és modellezési eszközök

A kutatás során a Python programozási nyelvet és a Jupyter Notebook környezetet használtam, mivel ezek kiválóan alkalmasak az adatmodellezésre, elemzésre és vizualizációra. A Python gazdag könyvtárkészlete lehetővé tette a rangsorolási algoritmus hatékony implementálását és a statisztikai kiértékelést is.

4.2. Használt adatkészlet és kiválasztott modellek

Az adatelemzéshez a statsbombpy adatkészletét használtam, innen kiválasztva a spanyol La Liga 2015/16-os szezon mérkőzéseinek adatait. Ez az adatkészlet tartalmazza a szezon összes csapatát és mérkőzésének eredményeit.

Az elemzés során három különböző Élő-alapú modellt teszteltem:

4.2.1. Baseline Élő modell

Ebben a modellben minden csapat egyenlő, 1500-as kezdőpontszámmal indul, így az előző szezon eredményei nem befolyásolják a kiindulópontot. A rangsor kizárólag az aktuális szezon mérkőzéseinek eredményei alapján frissül, és nem alkalmaz további dinamikus paramétereket, mint a hazai pálya előnye (HFA) vagy dinamikus döntetlenkezelés.

4.2.2. Uniform dinamikus Élő modell

Az Uniform modellben minden csapat ugyancsak 1500 pontról indul, de az aktuális szezon mérkőzései alapján frissített pontszámítási algoritmus dinamikus paramétereket is alkalmaz. A modell figyelembe veszi a hazai pálya előnyét (HFA), a döntetlen dinamikus kezelését, valamint egy K-faktort, amelyet az adott meccs gólkülönbsége állít be; mindezekkel a pontszámok módosulásának mértékét szabályozva.

4.2.3. Végso dinamikus Élő modell

Ebben a modellben a csapatok kezdő pontszámai az előző szezon teljesítményeire épülnek, így tükrözve az előző idény rangsorát. Az Uniform modellhez hasonlóan ez is dinamikus paraméterekkel dolgozik, vagyis HFA-t, döntetlen valószínűséget és egy K-faktort alkalmaz, lehetővé téve a szezon közbeni teljesítményváltozások érzékenyebb követését.

4.3. Paraméterek és egyedi beállítások

A dinamikus modell kidolgozása során az alábbi főbb paramétereket és beállításokat alkalmaztam:

- **K-faktor:** A mérkőzések utáni Élő pontok mozgását szabályozza. Alapértékként 20-at használtunk, amely a pontszámváltozás alapérzékenységet határozza meg. Ezt az értéket a gólkülönbségtől függően módosítottuk, hogy a jelentős különbséggel megnyert mérkőzések nagyobb hatást gyakoroljanak a rangsorra. Például, egy bajnoki mérkőzés esetében egy 3 gólos győzelem esetén az alapérték 35-re emelkedik, ami a győzelem mértékét tükrözi.
- **Hazai pálya előnye (HFA):** A HFA alapértéke 0,1, amely a hazai csapatnak biztosított pontkorrekció. A pontosabb szimuláció érdekében ez a különbség függ a csapatok közötti erőkülönbségtől, egy 0,0075-ös súlyozóval növelve a HFA-t nagyobb erőkülönbség esetén. Például, ha a hazai csapat pontszáma 200 ponttal magasabb az ellenfelénél, a hazai előny 0,1-ről 0,25-re emelkedik.
- **Dinamikus döntetlen valószínűség (Kappa):** A döntetlenek modellezésére egy dinamikus kappa-értéket vezettem be. Az alap kappa-érték 0,05, de a gyengébb csapatok nagyobb kappa-szorozót kapnak, ami növeli a döntetlenért járó pontnyereségüket. Ez a paraméter 0,025 és 0,1 között változhat a csapatok relatív erőssége alapján, így kiegyensúlyozottabbá téve a mérkőzés eredményét, különösen, ha erős és gyenge csapatok találkoznak.

4.4. Statisztikai elemzés

A rangsorok objektív összehasonlításához statisztikai rangkorrelációs módszereket használtam, melyekkel számszerűen mértem a modellek teljesítményét:

- **Spearman-féle rangkorreláció:** Ezt a módszert alkalmaztam a csapatok végső rangsorai közötti monoton kapcsolat mérésére, amely segít feltárni, mennyire követik a különböző modellek a klasszikus rangsor struktúráját. Ennek magas értéke azt jelezni, hogy az adott modell képes szorosan leképezni a hagyományos szezonbeli rangsort, azaz jól követi a teljesítmények relatív sorrendjét. [12]

$$r_s = 1 - \frac{6 * \sum d_i^2}{N * (N^2 - 1)} \quad (22)$$

Ahol:

- d_i : az egyes elemek közötti rangeltérés,
- N: a rangsorolt elemek száma.
- **Kendall Tau érték:** Ez a rangkorrelációs mérőszám a végső rangsorok közötti hasonlóságra és a pozíciók közötti közvetlen eltérésekre érzékeny. Ezzel a módszerrel részletesebben elemezhettem a csapatok rangsorának eltéréseit, így mélyebb betekintést nyerhettem abba, hogy a modellek mennyire térnek el az eredeti sorrendtől a helycserék számát tekintve. [13]

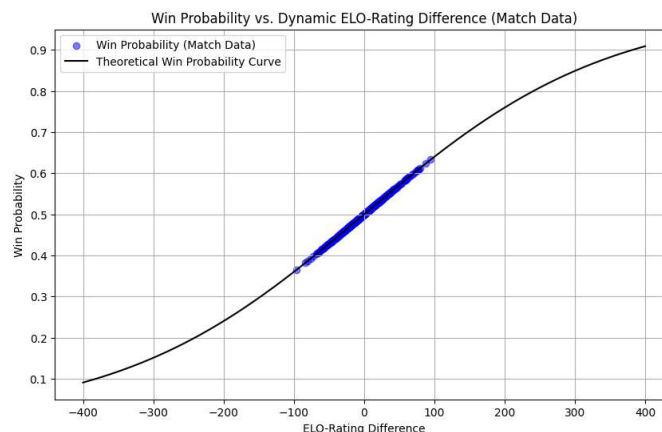
$$\tau_{Kendall} = \frac{C-D}{n*(n-1)/2} \quad (23)$$

Ahol:

- C: Azok a párok, ahol az egyik rangsorban egy adott elem alatti helyeken nagyobb rangok vannak.
- D: Azok a párok, ahol az egyik rangsorban egy adott elem alatti helyeken kisebb rangok vannak.
- n: A rangsorolt elemek száma.
- **Saját tau metrika:** Ez a metrika százalékos értékben fejezi ki, hogy a modell mennyire tér el a hagyományos rangsortól

$$\text{Tau (\%)} = 100 - \left(\frac{\text{Összesített Rangsortbeli Különbség}}{\text{Maximális Rangsortbeli Különbség}} * 100 \right) \quad (24)$$

A győzelmi valószínűség bemutatása pontkülönbség függvényében:



2. ábra: A mérkőzésekre vonatkozó győzelmi valószínűség és pontkülönbség közötti korreláció

Forrás: Saját szerkesztés

5. EREDMÉNY

5.1. Rangsorolási modellek eredményei

A három modell, valamint a klasszikus rangsor összehasonlítását végzem el, hogy rávilágítsak az eredményekben mutatkozó eltérésekre. Az összehasonlítás eredményei egyértelművé teszik, hogy az egyes megközelítések eltérően súlyozzák a csapatok közötti teljesítménybeli különbségeket, ami érdekes változásokat eredményez a csapatok végső helyezéseiben a klasszikus rangsorhoz képest.

5.2. Statisztikai korrelációk és Tau-metrikák értelmezése

Egy sajátos Tau-metrika, Spearman-féle rangkorreláció és a Kendall Tau segítségével mértem a csapatok közötti erőviszonyok és a rangsor változásait, így különböző aspektusokból ismerhetjük meg a szezonbeli sportszakmai forma hatásait.

A Baseline modell esetében mindhárom korrelációs mutató a legmagasabb volt: a saját Tau-metrika 88,42%, a Spearman-féle rangkorreláció 96,46%, és a Kendall Tau 88,30%. Ezek a magas értékek arra utalnak, hogy a Baseline szorosan követi a hagyományos rangsort, illetve, hogy kevésbé érzékeny a szezon közbeni teljesítményingadozásokra.

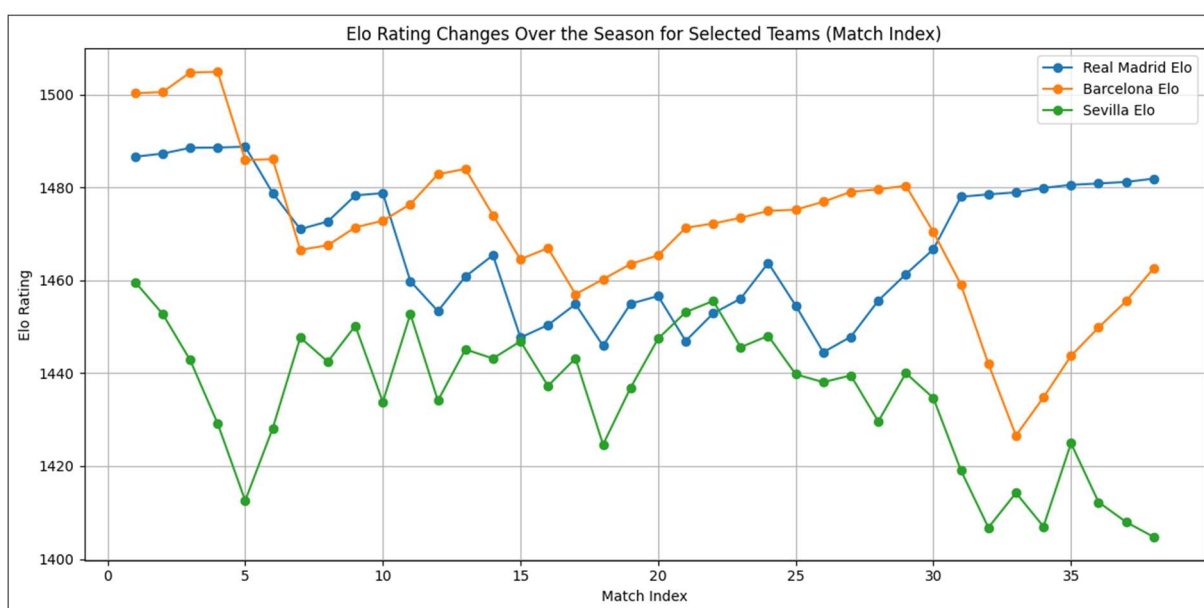
Az Uniform modell saját Tau-metrikája 52,11%, Spearman korrelációja 46,16%, és Kendall Tau értéke 36,17% – mind alacsonyabbak, mint a Baseline modellé, ami arra utal, hogy az Uniform modell kevésbé követi a hagyományos rangsort. Az alacsonyabb korrelációs értékek azt mutatják, hogy az Uniform modell jobban reagál az aktuális formaingadozásokra és képes eltérő rangsort eredményezni, mivel nincs kezdeti előny vagy hátrány a csapatok között.

A végső dinamikus modell minden metrikában hasonló eredményeket mutatott: 53,16%-os saját Tau-metrikát, 45,48%-os Spearman-féle rangkorrelációt, és 36,17%-os Kendall Tau értéket ért el. Ezek az értékek arra utalnak, hogy a modell jóval eltér a hagyományos rangsortól, különösen érzékeny a szezonbeli teljesítményváltozásokra és jól tükrözi az adott csapat aktuális jó vagy rossz szériáját.

La Liga 2015/2016 Season - Classic Rankings and Elo Ratings										
	team	points	GF	GA	W	D	L	Avg_Goals	Baseline Elo	Final Elo
1	Barcelona	91	112	29	29	4	5	2.95	1641	1462
2	Real Madrid	90	110	34	28	6	4	2.89	1660	1481
3	Atlético Madrid	88	63	18	28	4	6	1.66	1637	1459
4	Villarreal	64	44	35	18	10	10	1.16	1530	1407
5	Athletic Club	62	58	45	18	8	12	1.53	1548	1440
6	Celta Vigo	60	51	59	17	9	12	1.34	1526	1434
7	Sevilla	52	51	50	14	10	14	1.34	1483	1404
8	Real Sociedad	48	45	48	13	9	16	1.18	1490	1435
9	Málaga	48	38	35	12	12	14	1.00	1495	1435
10	Real Betis	45	34	52	11	12	15	0.89	1477	1428
11	Valencia	44	46	48	11	11	16	1.21	1460	1412
12	Las Palmas	44	45	53	12	8	18	1.18	1471	1420
13	Eibar	43	49	61	11	10	17	1.29	1438	1404
14	Espanyol	43	40	74	12	7	19	1.05	1462	1429
15	RC Deportivo La Coruña	42	45	61	8	18	12	1.18	1461	1419
16	Granada	39	46	69	10	9	19	1.21	1459	1436
17	Sporting Gijón	39	40	62	10	9	19	1.05	1455	1434
18	Rayo Vallecano	38	52	73	9	11	18	1.37	1450	1428
19	Getafe	36	37	67	9	9	20	0.97	1429	1414
20	Levante UD	32	37	70	8	8	22	0.97	1417	1419

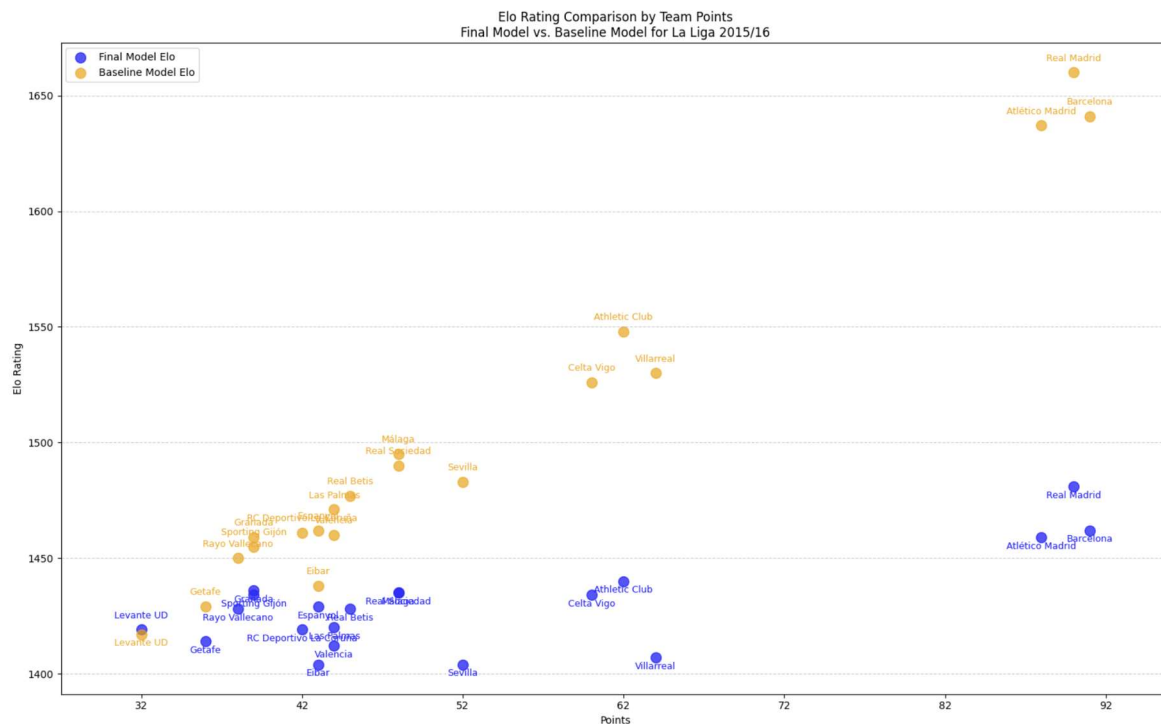
3. ábra: Klasszikus rangsor és Élő pontszámok összehasonlítása - La Liga 2015/2016 szezon

Forrás: Saját szerkesztés



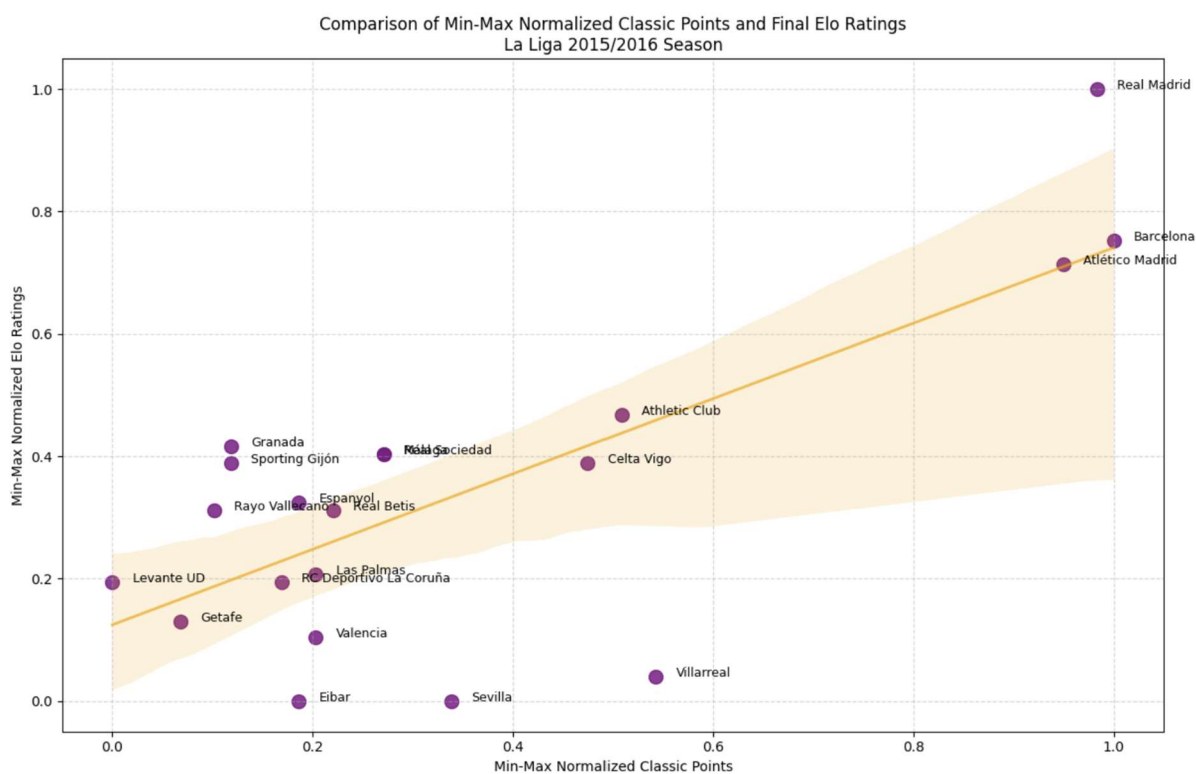
4. ábra: A Real Madrid, Barcelona és Sevilla Élő pontszám változása a 2015/2016-os szezon mérkőzése során

Forrás: Saját szerkesztés



5. ábra: A végső és az alap Élő pontszámok összehasonlítása – La Liga 2015/16 szezon

Forrás: Saját szerkesztés



6. ábra: A klasszikus ranglista és a végső dinamikus Élő modell összehasonlítása

Forrás: Saját szerkesztés

5.3. Győzelmi arányok az Élő pontkülönbség alapján

Az elemzés során megvizsgáltam, hogy egy csapat győzelmi aránya hogyan változik az ellenfelek Élő pontkülönbségének függvényében. Három kategóriát különítettem el:

- <25 pontos különbség: Ebben az esetben a két csapat pontértéke nagyon közel áll egymáshoz, tehát a küzdelem kiegyensúlyozottnak tekinthető.
- 25-50 pontos különbség: Itt már egy mérsékeltbb különbség mutatkozik a pontok között, amely az erősebb csapat mérsékeltbb előnyét jelzi.
- >50 pontos különbség: A nagy pontkülönbség arra utal, hogy az egyik csapatnak jelentős fölénye van.

Az eredmények azt igazolják, hogy minél nagyobb a pontkülönbség, annál nagyobb valószínűséggel nyeri meg az erősebb csapat a mérkőzést. A 2015/16-os La Liga szezont elemezve, azt az eredményt kaptam, hogy az átlagos győzelmi arány a <25 pontos különbségnél 40,71%, a 25-50 pont közötti kategóriában 32,89%, míg az 50+ pontos különbségnél 18,52%. Az eredmények azt jelzik, hogy a modell képes az Élő pontokból származó erőssorrend alapján meghatározni az esélyeket, és általánosságban az erősebb csapatok nagyobb valószínűséggel nyernek.

5.4. Becslés a mérkőzések kimenetelére

A dinamikus modell alapján létrehoztam egy predikciós táblázatot, amely minden mérkőzésre kiszámította a hazai és a vendég csapat győzelmi valószínűségét, a vizsgált szezon során. A predikciók csak a győztesek meghatározására koncentrálnak, a döntetlenekkel itt nem számolok. Az elemzés alapján a modell kimenetei az alábbi teljesítményjellemzőket mutatják:

- Helyes előrejelzések száma: A modell összesen 288 mérkőzés győztesére adott előrejelzést, amelyek közül 179 esetben helyesen jósolta meg a győztest. Ez 62,15%-os előrejelzési pontosságot jelent.
- Magasabb Élő-különbséggel rendelkező mérkőzések (pontkülönbség > 30): Azoknál a mérkőzéseknél, ahol a különbség nagyobb volt, a modell még pontosabbnak bizonyult. 89 ilyen mérkőzésből 70-et sikeresen előrejelzett, ami **78,65%-os pontosságot** jelent.
- A modell döntetlenre vonatkozó predikciói összhangban voltak a szezon valós eseményeivel. A döntetlen valószínűségek átlagosan 30%-os értéket mutattak, amely tükrözi a szezon tényleges eredményeit: a 380 mérkőzésből 92 végződött döntetlennel.

Az eredmények azt mutatják, hogy a modell képes az pontkülönbségeken alapuló erőviszonyok pontos leképezésére, és különösen megbízható előrejelzéseket ad azokban az esetekben, amikor az egyik csapat egyértelműen erősebb az értékelés alapján.

6. KITEKINTÉS ÉS KONKLÚZIÓ

6.1. A modell továbbfejlesztése és finomítása

A komplexitás különféle tényezők beépítésével növelhető. Egy lehetséges irány az egyes csapatok teljesítményére ható további adatpontok integrálása, mint például a játékosok aktuális formája, kulcsjátékosok sérülései, illetve akár olyan környezeti hatások, mint a mérkőzések időjárási körülményei. Ezek a változók még inkább finomítanák a számításokat, és lehetővé tennék, hogy a rangsorok még pontosabban tükrözzék az aktuális állapotokat.

6.2. Felhasználási lehetőség

A modell különösen jól alkalmazható lehet sportfogadások terén, ahol az aktuális forma pontos figyelembevétele jelentős előnyt nyújthat. A hagyományos rangsorolással szemben a dinamikus Élő-rangsor érzékenyen reagál a szezon közbeni formaingadozásokra, így a csapatok aktuális erőviszonyai jobban tükrözhetők. Ez sportfogadási szempontból azt jelenti, hogy a szorzók kialakításánál és a lehetséges kimenetek valószínűségének becslésénél pontosabb alapot adhat. A modell segítségével az olyan csapatok is pontosabban értékelhetők, akik a szezon során hullámzó formát mutatnak.

6.3. Konklúzió

A kutatás során kidolgozott dinamikus Élő modell jelentős előrelépést mutatott a csapatok szezonbeli teljesítményének pontosabb tükrözésében, a hagyományos rangsorolási rendszerekhez képest. Az eredmények azt mutatják, hogy a modellben alkalmazott dinamikus paraméterek – például a döntetlenek dinamikus kezelése és a hazai pálya előnye – finomítják a rendszert, így az érzékenyebbé válik a labdarúgás sajátos mérkőzés- és teljesítménydinamikájára. Az összehasonlítások egyértelművé tették, hogy a szezon során alkalmazott új rangsorolási szisztéma jobban közelíti a valós szezonbeli helyezéseket, mint a statikus megközelítések, és sikeres alkalmazását más sportágakban vagy sportfogadási rendszerekben is érdemes lehet vizsgálni.

Összességében a dinamikus rangsorolási rendszer nemcsak a csapatok közötti erőviszonyok pontosabb nyomon követésére alkalmas, hanem megalapozhatja a csapatok erőforrás- és stratégiai tervezését is, különösen a szezon közbeni teljesítményváltozások hatékony figyelembevételével.

7. IRODALOMJEGYZÉK

[1] Lašek, J., Szilávik, Z., & Bhulai, S. (2013).

The predictive power of ranking systems in association football. *International Journal of Applied Pattern Recognition*, 1(1), 27.

<https://doi.org/10.1504/ijapr.2013.052339>

[2] Hvattum, L. M., & Arntzen, H. (2010).

Using ELO ratings for match result prediction in association football. *International Journal of forecasting*, 26(3), 460-470.

<https://doi.org/10.1016/j.ijforecast.2009.10.002>

[3] Beggs, C., Shepherd, S., Emmonds, S., & Jones, B. (2017).

A novel application of pagerank and user preference algorithms for assessing the relative performance of track athletes in competition. *Plos One*, 12(6), e0178458.

<https://doi.org/10.1371/journal.pone.0178458>

[4] Neumann, C.; Duboscq, J.; Dubuc, C.; Ginting, A.; Irwan, A.M.; Agil, M.; Widdig, A.; Engelhardt, A.(2011)

Assessing dominance hierarchies: Validation and advantages of progressive evaluation with Elo-rating.

<https://doi.org/10.1016/j.anbehav.2011.07.016>

[5] Röhlcke, S., Bäcklund, C., Sörman, D., & Jonsson, B. (2018).

Time on task matters most in video game expertise. *Plos One*, 13(10), e0206555.

<https://doi.org/10.1371/journal.pone.0206555>

[6] Sánchez-Tójar, A., Schroeder, J., Farine, D. R., & McElreath, R. (2017).

Interaction frequency, network position and the role of competition in dominance hierarchy dynamics. *Oikos*, 126(6), 799-811.

<https://doi.org/10.1111/oik.04222>

[7] Routh, J., Paramasivam, S., Cockcroft, P., Wood, S., Remnant, J., Westermann, C., ... & Jeevaratnam, K. (2023).

Rating and ranking preparedness characteristics important for veterinary workplace clinical training: a novel application of pairwise comparisons and the Elo algorithm. *Frontiers in Medicine*, 10.

<https://doi.org/10.3389/fmed.2023.1128058>

[8] Szczecinski, L., & Djebbi, A. (2020).

Understanding draws in Elo rating algorithm. *Journal of Quantitative Analysis in Sports*, 16(3), 211-220.

https://www.researchgate.net/publication/341384358_Understanding_Draws_in_Elo_Rating_Algorithm

[9] García, M. S., Aguilar, Ó. G., & Romero, J. J. F. (2014).

Inconsistencies of the evaluation of home advantage in sports competitions under the three points per victory system. *Journal of human kinetics*, 42, 5.

https://johk.pl/wp-content/uploads/2023/03/10.2478_hukin-2014-0055.pdf

[10] Criado, R., García, E., Pedroche, F., & Romance, M. (2013).

A new method for comparing rankings through complex networks: model and analysis of competitiveness of major european soccer leagues. *Chaos an Interdisciplinary Journal of Nonlinear Science*, 23(4).

<https://doi.org/10.1063/1.4826446>

[11] Szczecinski, Leszek.

"G-Elo: generalization of the Elo algorithm by modeling the discretized margin of victory" *Journal of Quantitative Analysis in Sports*, vol. 18, no. 1, 2022, pp. 1-14.

<https://doi.org/10.1515/jqas-2020-0115>

[12] Xiao, C., Ye, J., Esteves, R. M., and Rong, C. (2016)

Using Spearman's correlation coefficients for exploratory data analysis on big dataset. *Concurrency Computat.: Pract. Exper.*, 28: 3866–3878.

<https://doi.org/10.1002/cpe.3745>

[13] Shiekh, Rega Hassan Ali and El-Hashash, Essam F. (2022)

A Comparison of the Pearson, Spearman Rank and Kendall Tau Correlation Coefficients Using Quantitative Variables. *Asian Journal of Probability and Statistics*. pp. 36-48.

<https://doi.org/10.9734/ajpas/2022/v20i3425>