



Budapesti Műszaki és Gazdaságtudományi Egyetem
Villamosmérnöki és Informatikai Kar
Távközlési és Médiainformatikai Tanszék

Hajósebesség-elemzés a SailGP-ben

OTDK dolgozat

Készítette:

Zentai Benedek

Konzulens:

Dr. Toka László

2024

Absztrakt

Az adatelemzés jelentősége a high-performance sportokban az utóbbi években jelentősen nőtt, lehetőséget teremtve a gépi tanulási technikák további alkalmazására. A SailGP egy viszonylag új versenysorozat a vitorlázásban, amelyet a sebesség és az innováció iránti szenvedély hajt, hasonlóan a Formula 1 versenyzésben tapasztalható fejlesztésekhez. Kutatásunk bemutatja az adatalapú megközelítések fontosságát a teljesítmény és a döntéshozatal javításában a legrangosabb vitorlaseseményeken. Adatbányászati technikák használatával egy 2021-ben Bermudán, illetve 2023-ban Olaszországban mért adathalmazt vizsgáltunk meg. A futamok adatainak elemzésével a tanulmány értékes betekintést nyújtott a változók közötti kapcsolatokba, mint például a szélesebbesség, a szélirány, a hydrofoil használata és a kormánylapát beállításai, valamint ezek hatására a hajósebességre. Különböző predikciós modelleket alkalmaztunk és értékeltünk: a Gradient Boosted, a Random Forest és egy kombinált modellt, különféle teljesítménymetrikák felhasználásával, mint például az R^2 érték és a négyzetes hiba. Az eredmények azt mutatják, hogy a modellek képesek a hajó sebességének pontos előrejelzésére. Ezek az eredmények felhasználhatóak a versenystratégiák finomításához, a vitorla és a kormánylapát beállításainak optimalizálásához és az összesített teljesítmény javításához a SailGP versenyeken. A jövőbeli terveink közé tartozik a SailGP-val való együttműködés nagyobb adathalmazokon, és a modellek integrálása valós idejű versenyzési helyzetekbe.

Tartalomjegyzék

1. Bevezetés	3
2. Háttér	3
3. Adathalmaz	4
4. Hajósebesség-elemzés	6
4.1. A kulcstényezők	7
4.2. Véletlen erdő modell	8
4.3. Gradient Boosted modell	9
4.4. Alapmodell teljesítmény	10
4.5. Együttes modell	12
5. Hajóemelkedés és értékelés	13
5.1. Eljárás	13
5.2. Feltáró elemzés	15
6. Következtetések	17

1. Bevezetés

A SailGP versenysorozat [11] egy izgalmas és forradalmi vitorlásbajnokságként robbant be a sportéletbe, amely a high-performance vitorlázás csúcsát mutatja be. A SailGP egyik igazán lenyűgöző aspektusa a hajók által elért sebesség és az ezzel járó adrenalin, ahogy a szél erejét kihasználva, a foilozás technológia segítségével a víz felszíne felett repülnek. A foilozás egy forradalmi újítás, amely lehetővé teszi a hajók számára, hogy rendkívüli sebességeket érjenek el és átlépjék azokat a határokat, amelyeket korábban a vitorlázásban felülmúlhatatlannak tartottak. A projekt célja, hogy rávilágítson a hajó sebessége és a különböző befolyásoló tényezők közötti összetett kapcsolatra, ezáltal hozzájárulva a high-performance vitorlázás szélesebb körű megismeréséhez és megértéséhez. A projektből származó eredmények és előrejelző modellek segíthetnek a vitorlázóknak, csapatoknak és taktikusoknak abban, hogy megalapozott döntéseket hozzanak és optimalizálják teljesítményüket a vízen. Ahogy a vitorlázás világa egyre inkább adatvezérelt megközelítéseket alkalmaz, ez a projekt mérföldkövet jelent az adatelemzés erejének kihasználása tekintetében, mely feltárja a SailGP és más magasrangú vitorlásversenyek sikerének titkait.

A vitorlás analitika területén végzett átfogó irodalom áttekintésünk során nyilvánvalóvá vált, hogy jelentős hiányosságok vannak a kifejezetten a SailGP-re fókuszáló kutatások terén. Annak ellenére, hogy egyre nagyobb érdeklődés mutatkozik az adatvezérelt megközelítések és analitikák iránt különböző sportokban, beleértve a vitorlázást is [1], kevés információ és tanulmány foglalkozik kifejezetten a SailGP kontextusában végzett vitorlás analitikával. Az irodalom hiánya arra utal, hogy még kiaknázatlan lehetőségek rejlenek az adatelemzés erejének felfedezésében és alkalmazásában a high-performance vitorlázás területén. A SailGP-hez kapcsolódó vitorlás analitikák kutatásának hiánya lehetőséget kínál új, innovatív megközelítések és módszertanok kifejlesztésére, amelyek kifejezetten ehhez az egyedi vitorlásversenyhez alkalmazkodnak.

Noha a SailGP csapatok által alkalmazott adatbányászati folyamatokról konkrét információk nem állnak rendelkezésre, feltételezhető, hogy az adatbányászat kulcsszerepet játszik teljesítményük javításában [4, 8, 3].

2. Háttér

A SailGP-vel összefüggésben több kulcsfontosságú terminológia is fontos szerepet játszik a sportág dinamikájának megértésében. A foilozás arra a technikára utal, amely során a katamarán testét a víz felszíne fölé emelik a hidrofoilok, melyek a testek alatt helyezkednek el. Ezek a hidrofoilok, amelyeket általában "uszony"-nak, vagy egyszerűen "foil"-nak hívnak, emelést és stabilitást biztosítanak a hajónak.

Ezenkívül a kormánylapátokon is lehetnek foilok, amelyek vízszintes uszonyokként helyezkednek el a testek hátulján. Ezek a kormányfoilok jelentősen hozzájárulnak a katamarán teljes emeléséhez és irányíthatóságához, növelve annak manőverezhetőségét, és sebességének maximalizálását a vízen.

Ami a vitorlákati illeti, a hajók innovatív szárnyvitorlákkal vannak felszerelve, amelyeket úgy terveztek, hogy a repülőgépek szárnyaira hasonlítsanak. Ezek a szárnyvitorlák kulcsfontosságúak a foilizás technikájában, mivel stabilitást és aerodinamikai hatékonyságot biztosítanak, hogy hatékonyan hasznosíthassák a szél erejét. A SailGP hajók innovatív kialakítása lehetővé teszi számukra, hogy szédítő sebességgel száguldjanak a víz felett. Ez a technológia, a vitorlázók tapasztalt navigációjával együttesen új szintre emeli a teljesítményt és a versenyképességet a vitorlázás világában.

3. Adathalmaz

A fejlett szenzorok és a különböző fedélzeti adatgyűjtő rendszerek használatával a csapatok hatalmas mennyiségű adatot gyűjtenek az edzések és a versenyek során. A SailGP adathalmaz, amelyet ebben a projektben a regressziós modellekhez használtunk, átfogó és részletes kimutatást ad egy-egy futamról, amely körülbelül 12 percet ölel fel. Ezt az adathalmazt különösen értékesé teszi a tény, hogy másodpercenként egy adatpontot tartalmaz, így pontos képet nyújt a verseny dinamikájáról és annak menetéről. A közel 60 rendelkezésre álló, a hajó sebességének, a szélesebességnek, a kormánylapátoknak, a hydrofoiloknak, a vitorláknak és más fontos alkatrészeknek a különböző állításait jegyző változókról, az adatkészlet számos numerikus adatot tartalmaz. Érdemes megjegyezni, hogy az adatok magas időbeli felbontással rendelkeznek.

Egy ilyen összetett adathalmaz beszerzése azonban jelentős kihívást jelentett a szigorú SailGP-adatok titkos kezelése miatt. Mivel az információk rendkívül bizalmasnak minősülnek, ezért az ilyen adathalmazhoz való hozzáféréshez komoly erőfeszítésekre és együttműködésre volt szükség. Az akadály leküzdése érdekében számos email kommunikációt kezdeményeztünk a SailGP-ben részt vevő vezető csapatokkal, segítségüket és együttműködésüket kérve a szükséges adatok megszerzéséhez. A folyamat magában foglalta a bizalom kiépítését, a projekt céljainak magyarázatát, valamint a csapatok biztosítását a legnagyobb titoktartásról és adatvédelemről. Kihívó kommunikáció és a csapatok együttműködése révén végül sikerült hozzáférni ehhez az értékes adatkészlethez. A SailGP adatszettben található adatok összetettsége és érzékenysége ritkaságnak számít a vitorlás analitika területén.

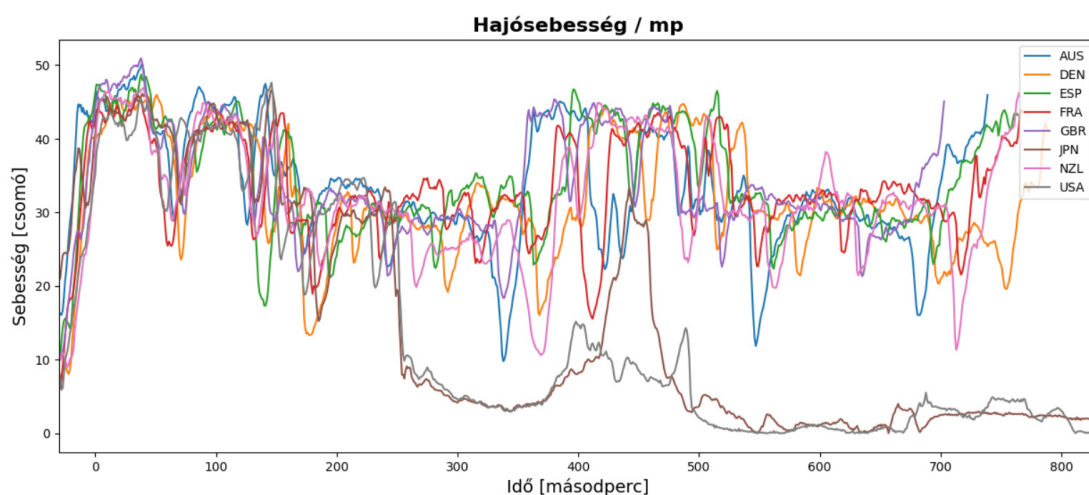
A különböző mértékegységek, mint például a szögek és csomók használata, tovább növeli az adatkészlet összetettségét. A változók sokfélesége és a magas időbeli felbontás egyedülálló lehetőséget kínál egy mélyreható betekintésre a versenydinamikába, a csapatok teljesítményébe és a különböző komponensek összhatására. Az adataik bizalmas kezelése és a titkosítottága jelentette kihívások sikeres leküzdésével, sikerült kihasználnunk a SailGP adatkészlet gazdagságát, ami részletes elemzések, modellezést és értékes betekintést nyújtott a high-performance vitorlázás rejtelmeibe.

Ez az adatkészlet a 2021-es bermudai esemény adatait tartalmazza, konkrétan a 4. futamét, amely a versenyhétvége második napján zajlott [9]. Ebben a futamban az összes hajó azért versenyzett, hogy biztosítsa a helyét a döntőbe jutásért, ahol a hétvégi verseny három legjobb csapata mérkőzött meg egymással.

Gyakornok adatelemzőként a SailGP-nél 2023 szeptemberében lehetőségem nyílt betekintést nyerni a professzionális vitorlázás mai világába. Ez az egyetlen hétig tartó, felbecsülhetetlen értékű tapasztalat nemcsak a sport iránti megbecsülésemet mélyítette el, hanem hozzáférést biztosított a legkülönbélebb adatkészletekhez is. A foilozás teljesítményének vizsgálata érdekében a tarantói (Olaszország) eseményről származó adathalmazt használtam fel, amely három különböző futam eredményeit foglalta magába. Ennek a kibővített adatkészletnek a beépítése jelentősen növelte elemzésünk átfogóságát, lehetővé téve a foilozási technikák és teljesítmény részletes vizsgálatát egy szerteágazó versenykörnyezetben.

4. Hajósebesség-elemzés

A hajó sebességét befolyásoló tényezők sokrétűek és egymással összefüggőek. A szél döntő szerepet játszik, mivel a csapatoknak stratégiaileg kell kihasználniuk az erejét a hajó előrehaladásához. A hatékony foilizási technikák, amelyek során a hydrofoilok segítségével a hajó kiemelkedik a vízből, jelentősen növelhetik a sebességet az ellenállás csökkentésével. A hajó optimális repülési magassága, amelyet a foilok precíz állításaival érnek el, kritikus a teljesítmény maximalizálása szempontjából. A hydrofoil és a kormánylapát beállítása közvetlenül befolyásolja a hajó stabilitását és manőverezhetőségét. Ezeknek az alkatrészeknek a finomhangolása lehetővé teszi a vitorlázók számára, hogy optimalizálják az emelkedést, az ellenállást és az irányítást, ezáltal hatékonyan tudjanak navigálni a kihívásokkal teli versenypályákon. A vitorlák, mint például a szárnyvitorla és az orrvitorla, kulcsfontosságúak a szélenergia befogásában, és előrehaladásra alakításában. A legénység pozíciója a fedélzeten szintén fontos szerepet játszik a hajó sebességében, mivel súlyelosztásuk befolyásolhatja a stabilitást és az egyensúlyt.



1. ábra. A grafikon a csapatok hajósebességét mutatja be a verseny teljes időtartama alatt. Az adatgyűjtés T-30 másodpercnél kezdődik, és az ábráról egyértelműen látszik, hogy minden csapat igyekszik maximalizálni a sebességét a rajt pillanatára. Megfigyelhető, hogy ezen a szakaszon és az első szakasz során vitorláznak a csapatok a legnagyobb sebességgel, ahol a GBR csapat a lenyűgöző 50,0 csomós (92,6 km/h) sebességet is eléri. Az ábrán az is látszik, hogy a manőverek során enyhe sebességcsökkenések tapasztalhatóak, és észrevehető sebességkülönbségek vannak a szél iránya felé és a hátszélben haladó szakaszok között. Érdeemes megjegyezni, hogy ebben a futamban a JPN és az USA csapata is ütközött az első széllel szembeni szakasz során, aminek következtében kiálltak a versenyből.

Ezeknek az összetevőknek az együttes elemzése hasznos betekintést nyújt abba, hogy a csapatok miként javíthatják a hajójuk sebességét. Az adatbányászati és gépi tanulási technikák alkalmazásával minták azonosíthatók és előrejelző modellek fejleszthetők a teljesítmény optimalizálása érdekében. Ez az elemzés lehetővé teszi a csapatok számára, hogy adatvezérelt döntéseket hozzanak a vitorla és a foil beállításokról, a legénység pozicionálásáról és az általános versenysztratégiáról, végső soron ezzel versenyelőnyt biztosítva számukra a SailGP gyors és dinamikus világában.

4.1. A kulcstényezők

A változók fontosságának elemzése értékes információkat nyújt arról, hogy a különböző komponensek milyen mértékben befolyásolják a hajó sebességét előrejelző regressziós modellünket. Random Forest és Gradient Boosting regresszormodelleket használtunk a változók jelentőségének kiszámításához, amelyet maguk az algoritmusok határoztak meg. Ezt követően a hajósebesség előrejelzésében betöltött relatív súlyaik alapján rangsoroltuk őket. Ezek az eredmények rávilágítanak minden egyes változó szerepére a modell előrejelző teljesítményében, lehetővé téve számunkra, hogy a legbefolyásosabb tényezőkre összpontosítsunk és azokat prioritásként kezeljük.

A vizsgált változók közül a legnagyobb hatással rendelkező komponens az `ANGLE_DB_RAKE_LIM_DEG` volt, amelynek súlya 0,696855. Ez azt jelzi, hogy a hydrofoilok dőlésszögének változásai befolyásolják a legjelentősebb mértékben a hajó sebességének előrejelzését. Ezt a szöget a foil hajótesthez viszonyított dőlésének mérésével határozzák meg. A mérést a hydrofoil első és hátsó élénél speciális műszerekkel és érzékelőkkel végzik. A dőlésszöget úgy állítják be, hogy a foilt előre és hátra mozgatják, illetve a foil hajótesthez viszonyított szögét módosítják. A foil dőlésszöge kulcsszerepet játszik a hajóra ható emelő- és légellenállási erők szabályozásában, ezáltal közvetlenül befolyásolja a hajó sebességét és manőverezhetőségét [6].

Hasonlóképpen, a kormánylapát szöge viszonylag magas súlyt kapott, 0,114463 értékkel. Ez azt sugallja, hogy a kormánylapát szögének változásai jelentősen hozzájárulnak a hajó sebességének ingadozásaihoz. A pontos és összehangolt kormánylapát-beállítások optimalizálhatják a hajó útvonalát és minimalizálhatják az ellenállást, ami a sebesség növekedéséhez vezet.

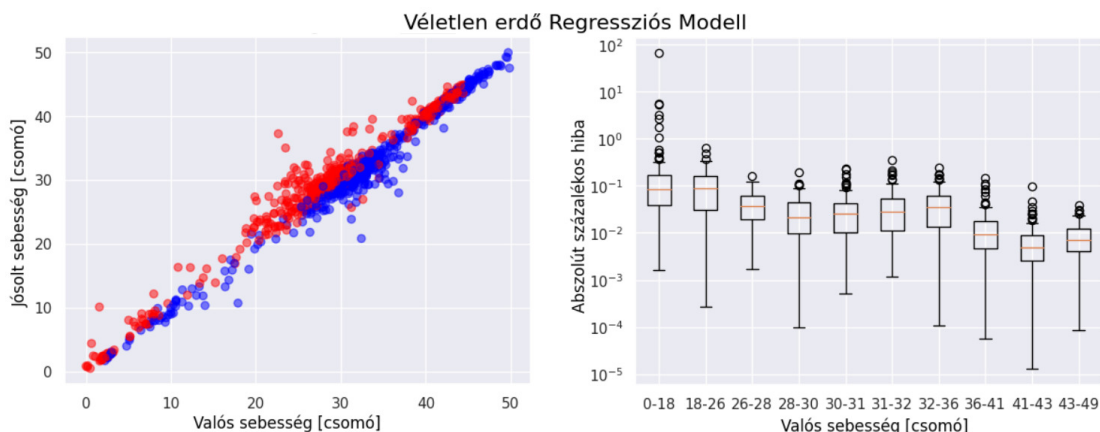
Másrészt, bizonyos jellemzők, amelyek a kormánylapát átlagos szögét és a foil billenésének mértékét jelzik egy adott helyzetben, kicsivel alacsonyabb súlyokat értek el: 0,034179 és 0,004883. A billenés (szakszóval: *cant*) arra utal, hogy a hydrofoil milyen szögben van oldalra billentve a hajó függőleges tengelyéhez képest. Bár ezek a változók egyenként kisebb hatással bírnak, mégis hozzájárulnak a modell összesített előrejelző képességéhez.

A kapott eredmények arra világítanak rá, hogy a hajótervezést és az irányítási mechanizmusok bizonyos aspektusait figyelembe kell venni a hajó sebességének pontos előrejelzéséhez. A foil, a platform hossza és a kormánylapát beállításaihoz kapcsolódó jellemzők kiemelt szerepe rámutat ezen elemek kiemelt jelentőségére a teljesítmény optimalizálásában. A végső regressziós modellt az irreleváns változók eltávolításával tökéletesítettük. Ezek az eredmények fontos alapul szolgálnak a hajó sebességét befolyásoló tényezők megértésében, valamint útmutatást adhatnak a jövőbeli teljesítménynövelő és optimalizáló erőfeszítésekhez a vitorlásversenyek, mint például a SailGP kontextusában.

4.2. Véletlen erdő modell

A véletlen erdő (angolul Random Forest) regresszió különösen alkalmas a nagy dimenziójú adatok kezelésére és a tényezők közötti összetett kapcsolatok rögzítésére [14]. Hatékonyan képes kezelni mind a kategóriális, mind a folytonos változókat, ami alkalmassá teszi a hajó sebességének előrejelzéséhez gyakran használt változatos adatkészletek elemzésére. Ez az együttes megközelítés több döntési fa kombinálásával javítja a modell robusztusságát és pontosságát.

A projekt során alkalmazott véletlen erdő regressziós modell ígéretes eredményeket hozott, igazolva hatékonyságát. Az eredményeket a TimeSeriesSplit nevű speciális keresztvalidációs módszerrel kaptuk, amely kifejezetten azokra az adatkészletekre van szabva, ahol az adatok időbeli sorrendje kulcsfontosságú [15]. A TimeSeriesSplit biztosítja, hogy a keresztvalidáció során a tanuló- és tesztalmozok megőrizték az adatok időbeli sorrendjét, elkerülve az adatszivárgást és megtartva az idősor struktúrájának integritását, valamint tükrözve a tesztkészleteken elért teljesítményt. A modell lenyűgöző R^2 értéket ért el, 0,94-et, ami azt jelzi, hogy a hajósebesség variáciájának körülbelül 94%-a magyarázható a kiválasztott változók alapján. Ez a magas R^2 érték erős korrelációt sugall a becsült hajósebesség és a valós értékek között.

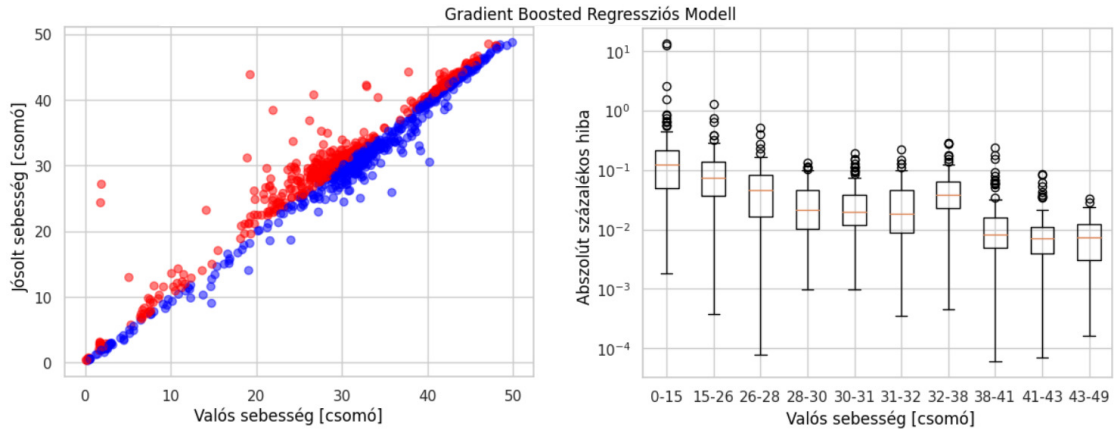


2. ábra. Ez a diagram vizuális összehasonlítást nyújt a tényleges és a modell által előrejelzett hajósebességek között. A diagramon a pontok eltérő színnel vannak jelölve: a piros pontok túlbecslést, míg a kék pontok helyes vagy alulbecslést jeleznek. A véletlen erdő regresszió diagramja a becült hajósebességek 30 csomó (55,6 km/h) körüli koncentrációját mutatja. Észrevehető, hogy a 25 csomó (46,3 km/h) és 30 csomó (55,6 km/h) közötti tartományban hajlamos túlbecsülni a sebességeket. Ez az észrevétel arra utal, hogy a modell ebben a konkrét tartományban kissé magasabb hajósebességeket jósol.

Ezeket túl, a véletlen erdő regressziós modell átlagos négyzetes hibáját (MSE) 4,5-re számítottuk. Az alacsonyabb MSE érték jobb modell teljesítményt jelez, mivel kisebb előrejelzési hibákat jelent. A 4,5-ös MSE arra utal, hogy az előrejelzett hajósebesség átlagosan körülbelül 2,1 csomóval tér el a valós értékektől.

4.3. Gradient Boosted modell

A Gradient Boosting regressziós modell alkalmazása a hajó sebességének előrejelzésére ebben a projektben számos előnyös tulajdonságon alapul, amelyek e modellezési technikát jellemzik. A Gradient Boosting a komplex kapcsolatok kezelésére és a nemlineáris mintázatok megragadására való képességéről ismert [13]. Ez különösen releváns a hajó sebességének előrejelzése szempontjából, mivel a sebességet befolyásoló tényezők bonyolult és nemlineáris dinamikát mutathatnak. A Gradient Boosting modellek robusztusak a túltanulás ellen, és hatékonyan képesek kezelni a magas dimenziójú jellemzőtereket. Az alkalmazott XGBoost modell kimagasló teljesítményt mutatott a hajó sebességének predikciójában. Az XGBoost modell a Boosting koncepcióján alapul, amely több gyenge tanuló, például döntési fák előrejelzéseit kombinálva hoz létre egy erős előrejelző modellt. A kiválasztott változók alapján az imponáló 0,95-ös R^2 értékkel a modell sikeresen magyarázza a hajó sebességének varianciáját. A Gradient Boosting regressziós modell 4,3-as átlagos négyzetes hibát



3. ábra. A modell alacsonyabb átlagos négyzetes hibát (MSE) mutat, ami a becsült és a valós hajósebességek közötti kisebb eltérést jelzi. Különösen figyelemre méltó, hogy a diagram a nagyobb sebességtartományokban javuló pontosságot mutat, különösen a 40 csomó (74 km/h) feletti hajósebességek esetében. Ez arra utal, hogy a modell előrejelzései szorosabban igazodnak a valós hajósebességekhez ezekben a magasabb tartományokban, ami nagyobb precizitást és pontosságot jelez ezeknél az predikcióknál.

(MSE) ért el, amely a becsült és a valós hajósebességek közötti átlagos négyzetes eltérést számszerűsíti.

4.4. Alapmodell teljesítmény

A viszonyítás (benchmarking) lehetővé teszi számunkra, hogy a modelleket egy előre meghatározott standard vagy referencia pont alapján értékeljük. Ebben az esetben az egyszerű bázismodell szolgált kiindulási alapként az összehasonlításhoz, amely egy optimalizálás nélküli, egyszerű megközelítést képvisel. Ez az egyszerű modell 0,92-es R^2 értéket ért el. A projekt során alkalmazott egyszerű bázismodell regressziós modellek használatát jelentette hiperparaméter-hangolás vagy modelloptimalizálás nélkül.

Ezzel szemben a mi modelljeink keresztvalidációt alkalmaztak a modell hiperparamétereinek szisztematikus finomhangolására, ezáltal optimalizálva az előrejelző képességeiket.

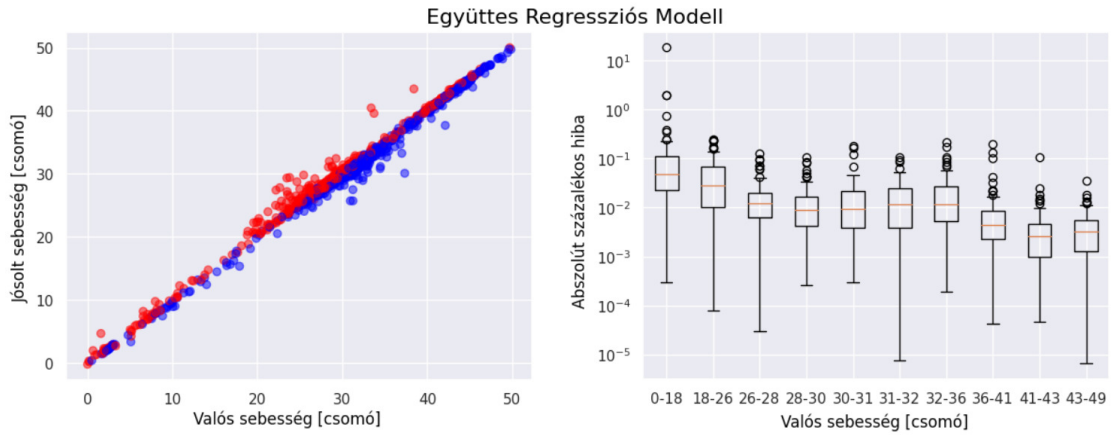
A hiperparaméter-hangolt modellek által elért magasabb R^2 értékek (0,94 és 0,95) azt jelzik, hogy ezek a modellek sokkal inkább képesek magyarázni a hajó sebességének változékonyságát a kiválasztott jellemzők alapján. Ez arra utal, hogy ezek a modellek jobban megragadják az adatokban rejlő mintázatokat és összefüggéseket, ami pontosabb előrejelzésekhez vezet. Az optimalizált hiperparaméterek lehetővé teszik a modell számára, hogy pontosabban alkalmazkodjon az adatkészlet finomságaihoz, mely javuló teljesítményt eredményez.

Emellett a k-szoros keresztvalidáció egy változatának alkalmazása modelljeinkben segít megelőzni a túltanulást azáltal, hogy a modell teljesítményét az adatok több részhalmazán értékeli. Ez a módszer biztosítja, hogy a modellt ne befolyásolja túlzottan egy adott tanuló-tesztelési felosztás, ezáltal javítva az általánosítási képességeit.

Összefoglalva, a TimeSeriesSplit-et alkalmazó regressziós modelljeink felülmúlták az egyszerű bázismodellt, magasabb R^2 értéket elérve. A hiperparaméterhangolási folyamat által biztosított optimalizálás és robusztus értékelés hozzájárult a modell teljesítményének javításához, így ez az eljárás lett a legjobb választás a hajó sebességének előrejelzésére ebben a projektben.

4.5. Együttes modell

A döntés, hogy egy összetett modellt hozzunk létre a véletlen erdő regressziós modell és a Gradient Boosting regressziós modell kombinálásával, azon alapult, hogy mindkét modell erősségeit kihasználjuk, és javítsuk az általános előrejelzési teljesítményt. Mindegyik modellnek megvannak a saját erősségei és gyengeségei az adatokban rejlő összetett kapcsolatok megragadásában. Az előrejelzéseik kombinálásával az összetett modell szélesebb körű mintázatokat képes feltárni és pontosabb előrejelzéseket adni. Az ötlet lényege, hogy az egyes modellek kimeneteiből tanuljunk, és egy meta-modellt építsünk, amely ötvözi azok erősségeit, végső soron a jobb predikciós teljesítmény elérése érdekében. A meta-modellt az átlagos négyzetes hiba (MSE) és az R^2 érték kiszámításával értékeljük, amelyek betekintést nyújtanak a modell pontosságába és illeszkedési precizitásába.



4. ábra. A diagram megmutatja a hajósebesség előrejelzésének hatékonyságát, különösen a 30 csomós (55,6 km/h) és magasabb sebességtartományokban. Ezenkívül az együttes modell kisebb eltérést mutat, ami arra utal, hogy pontos predikciókat képes adni csökkentett hibával. Érdeemes azonban megjegyezni, hogy mindegyik modell hajlamos a 30 csomó (55,6 km/h) körüli sebességek túlbecslésére, ami egy potenciális területet jelez a további finomítás számára.

Az elért eredmények fontos információkat tárnak fel a meta-modell hatékonyságáról. A kiszámított MSE érték 3,2, ami a becült és a valós hajósebességek közötti átlagos négyzetes eltérést jelzi. Az alacsonyabb MSE értékek jobb előrejelzési pontosságot sugallnak, és ebben az esetben az elért MSE viszonylag kis előrejelzési hibákat jelent. Továbbá, a 0,97-es R^2 érték azt mutatja, hogy a meta-modell mennyire képes magyarázni a hajósebességek varianciáját. A 1-hez közeli magas R^2 érték erős kapcsolatot jelez a becült és a valós értékek között. Ebben az esetben a 0,97-es R^2 érték arra utal, hogy a meta-modell jelentős mértékben megragadja a hajósebesség változékonyságát.

5. Hajóemelkedés és értékelés

A SailGP vitorlásversenyek összetett világának feltárására irányuló folyamatos törekvésünk során figyelmünk egy döntő tényezőre összpontosult, amely a versenyelőny alapvető elemévé vált: a foilozásra. Ez a "repülési" képesség teljesen új dimenziót hoz a sportba, izgalmas lehetőségeket és komoly kihívásokat kínálva. Ezért elengedhetetlen volt, hogy alaposan értékeljük a foilozási manőverek stabilitását befolyásoló tényezőket [10, 2].

Központi szerepet játszottak ebben a vizsgálatban a dinamikus dőlési (HEEL degree) és bólintási szögek (PITCH degree), amelyek a vitorlázás alapvető mutatói közé tartoznak. A dőlési szög azt mutatja meg, hogy a hajó mennyire dől oldalra, ami alapvető fontosságú a szél erejének kihasználásában és az irány megtartásában. A túlzott dőlés felboruláshoz vezethet, míg a túl csekély dőlés korlátozhatja a hajó teljesítményét. Hasonlóképpen, a bólintási szög a hajó előre- vagy hátrahajlását méri, amely lényeges tényező a kormányzás, az iránytartás és a merülés megakadályozásának szempontjából (amikor a hajó orra a víz felszíne alá merül). Ezek a szögek együtt alkotják azt a kényes egyensúlyt, amely meghatározza a vitorlás stabilitását, sebességét és végső soron versenyképességét.

Annak érdekében, hogy lehetővé tegyünk alapos vizsgálatunkat, egyedi hozzáférést kaptunk a 2023-as Tarantói SailGP esemény értékes adataihoz. Ez a versenyhétvége ideális alapot nyújtott a foilozás stabilitásának vizsgálatához. Kutatásunkat egy központi kérdés vezérelte: mikor állíthatjuk pontosan, hogy egy hajó repülési állapota stabilnak tekinthető?

5.1. Eljárás

A `find_peaks` [12] függvényt alkalmaztuk a high-performance F50 katamaránok dőlési és bólintási szögei szélsőértékeinek azonosításához. Az adatok csúcsainak megtalálása számos tudományos és mérnöki alkalmazásban gyakori feladat, például a fontos jellemzők azonosításában, vagy akár bizonyos gépi tanulási feladatokban. Célunk az volt, hogy alaposan megvizsgáljuk ezeket a csúcsokat, és körülbelül 5%-ukat szelektíven kizárjuk, hogy meghatározzuk a stabilitás standardját. Következésképpen, hosszadalmas megfontolás után a -5 és $+5$ fok közötti dőlési szögműködési értékek alkalmazása mellett döntöttünk, ami tükrözi a vitorlázás művészete és a stabilitás tudománya közötti egyensúlyt. Hasonlóképpen, a bólintási szögek esetében alapos elemzéseket végeztünk, és megállapítottuk, hogy a $-3,5$ és $+1,5$ fok közötti küszöbértékek használata biztosítja a legpontosabb ábrázolást. Ennek az aszimmetrikus döntésnek az oka az optimális repülés szükségességéből fakad, mivel a hajó orrának

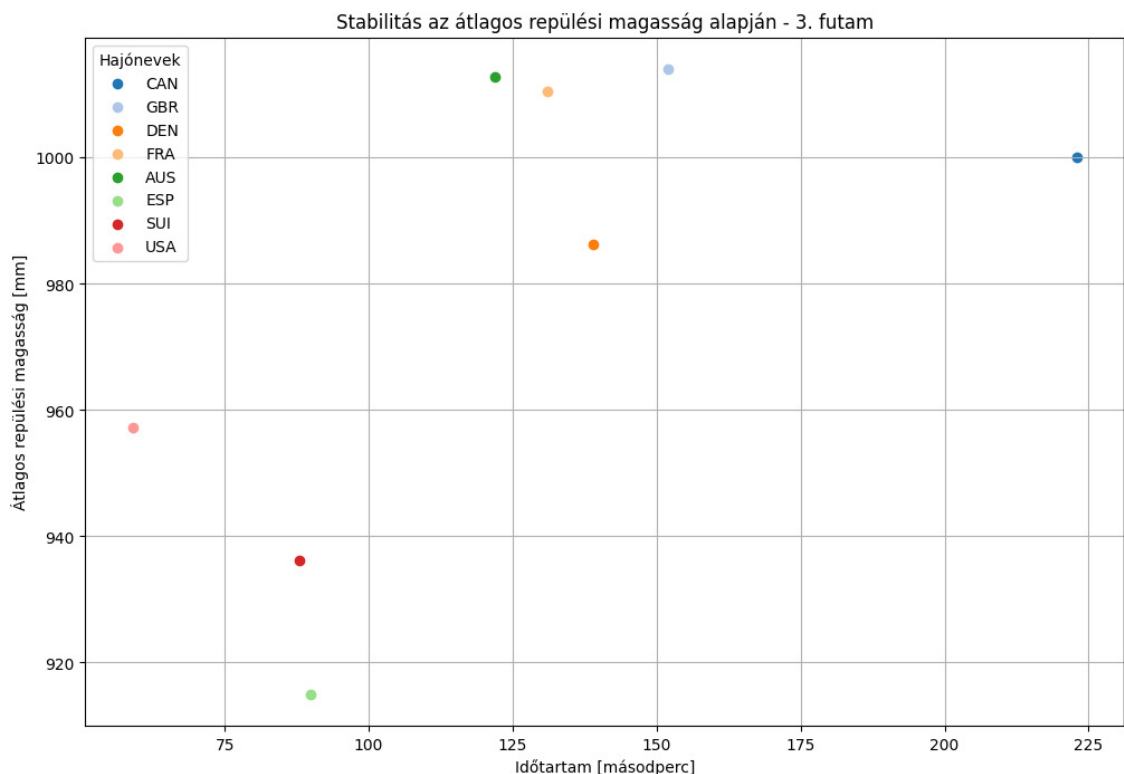
enyhe lefelé irányuló szögének fenntartása biztosítja, hogy a hajó a víz felszíne felett maradjon anélkül, hogy túlságosan érintené azt.

A felismerések iránti törekvésünk azonban nem ért véget ezeknek a küszöbértékeknek a meghatározásával. Továbbfolytattuk kutatásunkat, hogy azonosítsuk azokat a leghosszabb folyamatos időintervallumokat, amelyek alatt a hajók teljesítik ezeket a szigorú stabilitási feltételeket. Az elemzésünk összetettségének növelése érdekében bevezettünk egy kritériumot is: a minimális repülési magasságot. A többtestű hajók foilozása esetében ezt kifejezetten a szél alatti oldalon mérik, amely a széllal ellentétes, vagy a szélről védett oldalt jelenti. Vitorlázáskor a szél alatti oldal ellentétben áll a szél felőli oldallal, amely szembenéz a széllal. A szél alatti oldal kulcsfontosságú az egyensúly és az irányítás szempontjából, különösen a szél erejének kihasználásakor. A foilozás tekintetében a szél alatti testen elért repülési magassághoz különösen fontos, hogy az legalább 0,5 méterrel a víz felszíne felett legyen. Ez biztosítja az optimális teljesítményt és minimalizálja az ellenállást. A hajó szél alatti oldalának azonosításához a valós szélirány (True Wind Angle, TWA) adatait használtuk, amelyeket az orrnál mértek.

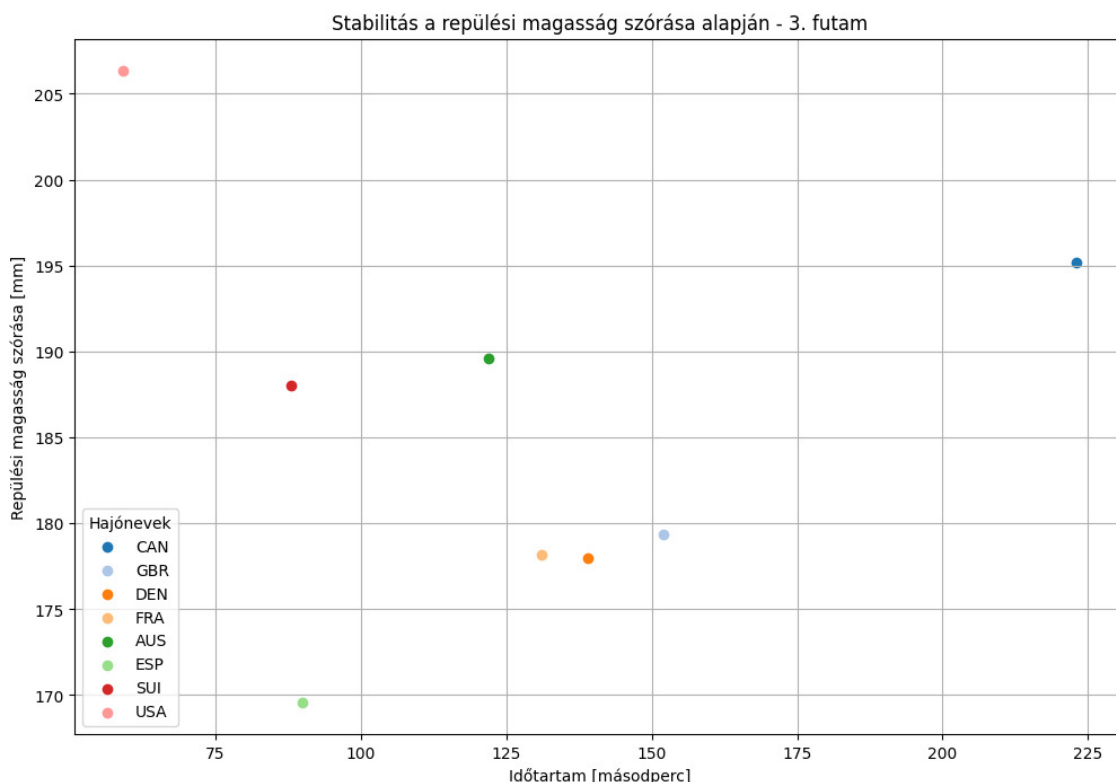
Kiszámítottuk az egyes hajók átlagos repülési magasságát azokban az időszakokban, amikor a legnagyobb stabilitást mutatták. Jól ismert tény, hogy minél magasabbra emelkedik a hajó a víz felszíne fölé, annál nagyobb kihívást jelent a stabil repülés fenntartása, és ezek a kihívások exponenciálisan növekednek. Ezért az átlagos repülési magasság elemzése ezekben az időszakokban alapos betekintésre adott lehetőséget a magasság és a stabilitás dinamikus kölcsönhatásába. Ezen felül részletes vizsgálatot végeztünk a repülési magasság szórására vonatkozóan, felismerve, hogy ez a paraméter elengedhetetlen a foilozás stabilitásában rejlő finom eltérések megértéséhez. Ez a dimenzió döntő szerepet játszott abban, hogy átfogóan megértsük, hogyan befolyásolják a repülési magasság ingadozásai a hajók általános stabilitási profilját.

5.2. Feltáró elemzés

Számításainkat gondosan dokumentálva olyan vizuális ábrázolásokat készítettünk, amelyek szorosan követték a tudományos kutatás alapelveit. Ezek a vizualizációk alaposan szemléltetik a hajók leghosszabb stabil időszakait és az ezekhez tartozó átlagos repülési magasságokat, árnyalt képet adva a tartósság és a magasság közötti bonyolult kapcsolatról. Ezzel párhuzamosan olyan grafikus illusztrációkat is készítettünk, amelyek bemutatják a leghosszabb stabil időszakok és a repülési magasság szórása közötti összefüggést. Ezek a vizuális elemek túlmutatnak az egyszerű statisztikákon, és érthető módon mutatják be kutatásunk gondos folyamatát.



5. ábra. A diagram bemutatja az átlagos repülési magasság és a leghosszabb időtartam közötti kapcsolatot, amikor a hajók teljesítik az összes feltételt a 3. futam során. Minden pont egy különálló hajót jelöl, a szín pedig a hajó nevét mutatja. A diagramról látható, hogy a DEN, a FRA és a GBR hajók kiemelkedőek a foilozás stabilitásában, magas átlagos repülési magasságot érnek el hosszú időtartamokkal kombinálva, ezzel a dobogóra kerülve. Emellett a CAN és az AUS hajók a 4. és 5. helyet szerezték meg.[7].



6. ábra. Az ábra szemlélteti a repülési magasság szórása és a leghosszabb időtartam közötti kapcsolatot a különböző hajók esetében a 3. futamban. Az ábra azt mutatja, hogy a DEN, a FRA és a GBR figyelemre méltó stabilitást tanúsítanak, alacsony repülési magasság szórással, ami következetes foilozási teljesítményt jelez. Ezzel szemben, bár a CAN rendelkezik a leghosszabb időtartammal, magas szórása fokozott instabilitást sugall, ami potenciálisan veszélyeztetheti a foilozás hatékonyságát.

Az adatok azt mutatják, hogy lehetséges összefüggés van a foilozás stabilitása és a verseny eredményei között. Figyelembe véve azt, hogy ezek a versenyek tipikusan 750-800 másodpercig tartanak, a stabil időtartamok hossza közel 150 másodperc—ami nagyjából a verseny teljes idejének 20%-a—ami növeli annak valószínűségét, hogy egy csapat dobogós helyezést ér el. Gyakran ezek a hajók csupán néhány másodperc különbséggel fejezik be a versenyt, így az eredmények felületes szemlélete nem feltétlenül tükrözi pontosan a foilozási teljesítményüket.

Ezek a betekintések különösen értékesek lehetnek a napi összefoglalók vagy a versenyek közötti jelentések elkészítésében, amelyeket a flight-controller számára készítenek. A flight-controller feladata az emelési és ellenállási erők finom egyensúlyának kezelése, biztosítva, hogy a hajó stabil maradjon, miközben a hydrofoiljain "repül" a víz felett. Ez a szerep egyedi technikai jártasságot és vitorlázási intuíciót igényel. Amikor a hajó felemelkedik a vízről, a flight-controller szakértelműen határozza meg a hajó repülési magasságát, szögét és stabilitását, közvetlenül befolyásolva a sebességet és az általános versenystratégiát [5]. A SailGP és más foilozási versenyek

nagy sebességű környezetében a flight-controller döntései jelenthetik a különbséget a győzelem és a vereség között. Ennek a szerepnek az evolúciója kiemeli a hagyományos vitorlázási készségek és a modern technológia egyesítését, hangsúlyozva a sport új korszakba való átmenetét. Figyelembe véve a szerep összetettségét, ezek a megállapítások jelentős mértékben hozzájárulhatnak egy csapat repülési képességeinek javításához egyik versenyről a másikra.

6. Következtetések

A SailGP az egyre gyorsabban fejlődő és rendkívül versenyképes vitorlázás élvonalában áll. Ez a dinamikus sportág csúcstechnológiájáról, intenzív versenyeiről és az adatvezérelt stratégiákra való erős támaszkodásáról ismert. Ahogy a SailGP továbbfejlődik és feszegeti a vitorlázás lehetőségeinek határait, az analitika és az adatbányászat szerepe egyre fontosabbá válik. Érdeemes megjegyezni, hogy kutatómunkánk a SailGP kontextusában úttörő jelentőségű az adatbányászat területének feltárásában. Az ismeretlen területre merészkedve célunk, hogy rávilágítsunk az adatok kiaknázásában rejlő lehetőségekre, és ezáltal teljesítménynövekedést érjünk el ebben az izgalmas sportágban.

Összességében az eredmények azt mutatják, hogy a RF és GBM modellek predikciói, valamint a meta-modellt magában foglaló összetett regressziós modell ígéretes eredményeket hoz a hajósebességek előrejelzésében. A viszonylag alacsony MSE és a magas R^2 érték arra utal, hogy a modell jelentős mértékben feltárja és magyarázza a hajósebességek varianciáját. Ezek az eredmények azt mutatják, hogy a modell hasznos a hajósebességek pontos becslésében és értékes a teljesítmény optimalizálásában a vitorlásversenyeken.

A foilozás teljesítményéről végzett elemzésünk továbbhangsúlyozza a foilozás stabilitása és a versenyeredmények közötti összetett kapcsolatot. Az általunk kidolgozott módszer alkalmas a foilozási teljesítmény értékelésére, és bármely többtestű foilos vitorlášhajón alkalmazható. Azok a csapatok, amelyek a verseny teljes idejének jelentős részében stabil foilozási időtartamokat mutatnak, gyakran nagyobb valószínűséggel érnek el dobogós helyezéseket. Ez a felfedezés továbbhangsúlyozza a flight-controller szerepének fontosságát és a foilozási metrikák folyamatos elemzésének potenciális előnyeit a versenystratégiák finomítása érdekében.

A csapatok a teljesítmények optimalizálására irányuló elemzése egyre inkább kiemeli az adatvezérelt döntéshozatal fontosságát a versenysportban. A SailGP-ben, ahol a győzelem és a vereség közötti különbségek gyakran minimálisak, az adatok hatékony felhasználása létfontosságú a versenyelőny megszerzéséhez.

Reményeink szerint izgalmas lehetőség rejlik a SailGP-val való együttműködésben, amely hozzáférést biztosít további adatkészletekhez és potenciálisan valós idejű versenyadatokhoz. A SailGP-val szoros partnerségben célunk, hogy előrejelző modelleinket szélesebb körű scenáriókra terjesszük ki, és pontosabb előrejelzéseket érjünk el élő versenyek során.

Hivatkozások

- [1] Flying across the sea, propelled by ai, 2021.
URL <https://www.mckinsey.com/capabilities/mckinsey-digital/how-we-help-clients/flying-across-the-sea-propelled-by-ai>. Utolsó elérés: 2024. május 24.
- [2] How sailgp's foiling f50 catamarans sail so much faster than the wind. URL <https://bitly.cx/xbSqo>. Utolsó elérés: 2024. május 24.
- [3] Oracle cloud infrastructure.
URL <https://www.oracle.com/customers/sailgp/>. Utolsó elérés: 2024. május 24.
- [4] Oracle help center, 2021. URL <https://docs.oracle.com/en/solutions/sailgp-on-oci/index.html>. Utolsó elérés: 2024. május 24.
- [5] Paul goodison on how to drive a foiling f50 in sailgp.
URL <https://www.yachtingworld.com/catamaran-sailing/paul-goodison-on-how-to-drive-a-foiling-f50-in-sailgp-132850>.
Utolsó elérés: 2024. május 24.
- [6] Jonathan Ridley: Sailgp: how do supercharged racing yachts go so fast? an engineer explains, 2019. szeptember. URL <http://bit.ly/40cLoB8>. Utolsó elérés: 2024. május 24.
- [7] Rockwool italy sail grand prix: Taranto - eredmények.
URL <https://bitly.cx/uoCH>. Utolsó elérés: 2024. május 24.
- [8] Sailgp.
URL <https://sailgp.com/news/nobu-katori-joins-japan/>. Utolsó elérés: 2024. május 24.
- [9] Sailgp bermuda eredmények, 2021.
URL <https://sailgp.com/news/bermuda-season-2-results/>. Utolsó elérés: 2024. május 24.

- [10] Sailgp f50 v america's cup ac50 catamarans. URL <https://yachtracing.life/sailgp-f50-v-americas-cup-ac50-catamarans/>. Utolsó elérés: 2024. május 24.
- [11] Sailgp kezdőlap. URL <https://sailgp.com/>. Utolsó elérés: 2024. május 24.
- [12] Scipy find_peaks. URL <https://bitly.cx/sqtyu>. Utolsó elérés: 2024. május 24.
- [13] Sklearn gradient boosting regression. URL <https://bit.ly/3YoA6ch>. Utolsó elérés: 2024. május 24.
- [14] Sklearn random forest regressor.
URL <https://bit.ly/3Kgkd1p>. Utolsó elérés: 2024. május 24.
- [15] Sklearn time series split. URL <https://bit.ly/3QdXYNv>. Utolsó elérés: 2024. május 24.