

INFORMATIKA I ANALITIKA U SPORTU

INFORMATICS AND ANALYTICS IN SPORTS

Valter Perinović¹, Matej Kumiša²

¹Tehničko veleučilište u Zagrebu, Vrbik 8, Zagreb, Hrvatska

²Tehničko veleučilište u Zagrebu, Vrbik 8, Zagreb, Hrvatska, Student

SAŽETAK

Rad opisuje što su informatika i analitika u sportu. Kako mogu pomoći razvoju sporta i poboljšanja performansi igrača i timova koristeći moderne tehnologije za prikupljanje informacija i matematičke jednadžbe za izračunavanje statistika.

Ključne riječi: *Informatika u sportu, Analitika u sportu*

ABSTRACT

The paper describes what are IT and analytics in sports and how they can help in development of sport and improve player and team performance using modern technologies for data acquiring and mathematical equations to calculate statistics.

Keywords: *IT in sport, Analytics in sport*

1. UVOD

1. INTRODUCTION

Analitika je bilježenje, tumačenje i obrada relevantnih statističkih podataka. Što više takvih podataka ima time će i sama statistika koja je produkt analitike biti točnija i korisnija.

Informatika je disciplina koja se bavi prijenosom, pohranom i manipulacijom podataka.

Informatika i analitika u sportu su međusobno usko vezane jer informatika nudi pohranu i prikupljanje statistički relevantnih podataka u ekipnim i pojedinačnim sportovima. Postoje određene baze podataka s raznim statistikama koje su skupljene informatičkim načinom unatrag više godina tada se analitičkim metodama obrađuju i donose zaključci.

Osim praćenja rezultata i statistika pojedinaca i ekipa informatika s napretkom tehnologije omogućuje razne mjere praćenja sportaša na samom treningu i time omogućava bolji uvid znanstvenicima u ljudsko tijelo i automatski novim saznanjima omogućava napredak samih trenera i time bolje fizičke performanse samih sportaša.

Samo prikupljanje podataka sa snimanjem većine većih natjecanja i pristupačnost kamerama pojedincima i klubovima znatno je olakšalo prikupljanje bitnih informacija. Analitika i informatika u sportu uglavnom imaju jasan cilj pružiti timu ili pojedincu bolji uvid u protivnike i u svoje mogućnosti te postići prednost nad istima.

2. ANALITIKA U SPORTU

2. ANALYTICS IN SPORTS

Analitika u sportu je jedno od najvećih napredaka u sportu proteklih godina. Korištenjem matematičkih modela i analiziranje velikog broja podataka koji su dostupni kako bi analizirali sposobnosti, trendove i uzorke u svrhu poboljšanja produktivnosti i predviđanja ishoda sportskih susreta.

Danas, svaki veći profesionalni sportski klub ima ili odjel za analitiku ili stručnjaka za analitiku. Timovi ljudi moraju skenirati bilješke od skauta, pretvoriti te PDF datoteke u Excel, a zatim ih prosljediti vrhunskim analitičarima. Nakon toga, te podatke matematičari trebaju analizirati i konvertirati u bitne statistike koje skauti i menadžeri koriste kako bi se utvrdilo koji igrač je najprikladniji i u tom trenutku najpotrebniji timu. Sve je dio stvaranja cjelokupnog profila igrača kako bi se utvrdilo je li igrač dobra investicija bilo potpisivanjem ugovora ili kupnjom od drugog kluba.

Analitika je sadašnjost i budućnost profesionalnog sporta. Svaki tim koji ih ne primjenjuje gubi kompetitivnost s ostalim timovima.

Popularnost donošenja odluka temeljenih na analitičkim podacima u sportu preneseno je na navijače koji konzumiraju više analitičkih sadržaja nego ikad. Sada postoje cijele web stranice posvećene istraživanju i analizi sportskih statistika i kako su povezane s predviđanjem rezultata. Primjer je FiveThirtyEight.com koji je započeo u ožujku 2008. godine Nate Silver. Web stranica, koja je u vlasništvu ESPN-a (Entertainment and Sports Programming Network) ima više od 20 novinara koji broje i pišu brojeve za navijače kako bi bolje razumjeli nadolazeću natjecateljsku sezonu. Silverove metode pokazale su se tako uspješne u sportu da ih je počeo primjenjivati na politiku. 2008. godine njegov je sustav analitike predvidio eventualni ishod kako će 49 od 50 država glasovati na predsjedničkim izborima te godine[1].

Koristeći sofisticirani sustav za obradu brojeva, web stranica kao što je FiveThirtyEight.com može procijeniti prethodne rezultate, omjer pobjede i poraza i povijest protivnika kako bi odredio ishod budućeg sportskog događaja. Prije nego što susret počne, fivethirtyeight.com može doći do najvjerojatnijeg rezultata na temelju brojeva za razliku od čistog instinkta na koji se do sad oslanjalo.

$$CCC = TCC (\textit{Home Team}) * TCC (\textit{Visiting Team}) * 1.27 \quad (2.1)$$

$$AGS = ((GS - OPP_DEF) / (\text{MAX}(0.25, OPP_DEF * 0.424 + 0.548))) * (AVG_BASE * 0.424 + 0.548) + AVG_BASE \quad (2.2)$$

$$AGA = ((GA - OPP_OFF) / (\text{MAX}(0.25, OPP_OFF * 0.424 + 0.548))) * (AVG_BASE * 0.424 + 0.548) + AVG_BASE \quad (2.3)$$

Sami izračun za na primjer nogomet se vrši pomoću SPI(soccer power index) koeficijenta. Nate Silver u svojem članku opisuje kako se SPI koeficijent može izračunati u 4 koraka[2].

U prvom koraku se računa CCC (combined competitiveness coefficient) prikazan formulom (2.1) gdje je TCC (team competitiveness coefficients) domaće ekipe pomnožen s TCC-om gostujuće ekipe i pomnožen konstantom. CCC se računa za svaku pojedinu utakmicu koju imaju u bazi podataka. TCC predstavlja koliko je igrača iz prve postave sudjelovalo u pojedinom susretu te se tada uvrštava u jednadžbu. U drugom koraku se računa AGS(adjusted goals scored) po izrazu (2.2) i AGA (adjusted goles allowed) po izrazu (2.3) kako bi se izračunali koeficijenti tima u napadu i obrani. U formuli su GS (goals scored) GA (goals allowed) i konstanta prosječan broj golova u internacionalnim utakmicama postignut po broju utakmica (AVG_BASE). U ovim formulama se računa u svakom pojedinom susretu „OFF“ i „DEF“ koeficijent za pojedinu ekipu, ali kao što se može primijetiti u tom susretu se u formulu uračunavaju i ti koeficijenti koji su pripadni protivničkoj ekipi (OPP_DEF, OPP_OFF) što u osnovi dovodi do matematičkog problema jer se u istom susretu i za protivnika računa po istim formulama. To dovodi do beskonačne petlje jer i ako bi se određivao koeficijent tima u napadu i obrani po prijašnjim utakmicama opet bi se samo vrtjelo u krug. To je riješeno putem iteracija u kojem daju početnu pretpostavku o kvaliteti protivnika (OPP_OFF, OPP_DEFF) za svaki susret u bazi podataka i onda pokreću program kroz nekoliko ponavljanja i svakim ponavljanjem dodaju nove informacije kako bi dobili stabilne i precizne koeficijente. Treći korak izračunavanja SPI koeficijenta je izračunavanje kvalitete svakog pojedinog igrača. U tom postupku se s obzirom na nastupe svakog pojedinog igrača u klubu, računa i dodjeljuje njegov napadački i obrambeni koeficijent. U četvrtom koraku se rezultati prethodnih koraka zbrajaju i tako tvore SPI koeficijent. Sami izračun za nogomet se vrši pomoću SPI(soccer power index) koeficijenta. Nate Silver u svojem članku opisuje kako se SPI koeficijent može izračunati u 4 koraka[2]. U prvom koraku se računa CCC (combined competitiveness coefficient) prikazan formulom (2.1) gdje je TCC (team competitiveness coefficients) domaće ekipe pomnožen s TCC-om gostujuće ekipe i pomnožen konstantom. CCC se računa za svaku pojedinu utakmicu koju imaju u bazi podataka. TCC predstavlja koliko je igrača iz prve postave sudjelovalo u pojedinom susretu te

se tada uvrštava u jednadžbu. U drugom koraku se računa AGS (adjusted goals scored) po izrazu (2.2) i AGA (adjusted goles allowed) po izrazu (2.3) kako bi se izračunali koeficijenti tima u napadu i obrani. U formuli su GS (goals scored) GA (goals allowed) i konstanta prosječan broj golova u internacionalnim utakmicama postignut po broju utakmica (AVG_BASE). U ovim formulama se računa u svakom pojedinom susretu „OFF“ i „DEF“ koeficijent za pojedinu ekipu, ali kao što se može primijetiti u tom susretu se u formulu uračunavaju i koeficijenti koji su pripadni protivničkoj ekipi (OPP_DEF, OPP_OFF). U osnovi to dovodi do matematičkog problema jer se u istom susretu i za protivnika računa istim formulama. To dovodi do beskonačne petlje jer i ako bi se određivao koeficijent tima u napadu i obrani po prijašnjim utakmicama opet bi se samo vrtjelo u krug. To je riješeno putem iteracija u kojem daju početnu pretpostavku o kvaliteti protivnika (OPP_OFF, OPP_DEFF) za svaki susret u bazi podataka. Pokreću program kroz nekoliko ponavljanja i svakim ponavljanjem dodaju nove informacije kako bi dobili stabilne i precizne koeficijente. Treći korak izračunavanja SPI koeficijent je izračunavanje kvalitete svakog pojedinog igrača. U tom postupku se s obzirom na nastupe svakog pojedinog igrača u klubu, računa i dodjeljuje njegov napadački i obrambeni koeficijent. U četvrtom koraku se rezultati prethodnih koraka zbrajaju i tako tvore SPI koeficijent.

Košarka je jedan od najboljih primjera kako je analitika promijenila način na koji se igra i mjeri izvedba igrača. NBA timovi sada koriste oblik tehnologije pod nazivom "Praćenje igrača", koja procjenjuje učinkovitost tima analizom pokreta igrača. Prema web stranici softvera SportVu, timovi u NBA-u sada koriste šest kamera instaliranih na stropu arena kako bi pratili pokrete svakog igrača na igralištu i košarkaške lopte 25 puta u sekundi. Prikupljeni podaci pružaju mnoštvo inovativnih statistika koje se temelje na brzini, udaljenosti, razdvajanju igrača i posjedu lopte. Neki primjeri uključuju kako se igrač brzo kreće, koliko je pretrčao tijekom igre, koliko je puta dotaknuo loptu, koliko je puta dodao, koliko je imao mogućnosti za nadoknadu i još mnogo toga. Dok tehnologije poput "Praćenje igrača" izgledaju kao val budućnosti, postoji razina težine

u određivanju kako iskoristiti višak informacija koje pruža kako bi igrači i treneri dobili prednost. Zbog velike količine podataka, koje se može dobiti, najveći problem je uočiti koji podaci su relevantni za razvoj tima i izbor igrača.

Sama količina podataka može biti velika. Dva teška čimbenika su kako će predviđanja podataka biti povezana sa stručnim osobljem u timu, napadačkim i obrambenim sustavom. Sve se to razlikuje od tima do tima. Sama analitika nije jamac timu da će biti vrhunski zato što bi po statistikama trebao biti. Na kraju, potrebno je pregledavanje faktora od strane profesionalnih trenera i stručnjaka za taktiku kako bi tim bio na svom vrhuncu.

Sportska analitika i dalje će se razvijati. Stručni stožeri timova će se više oslanjati na njih, ali još uvijek postoje načini poboljšanja.

3. INFORMATIKA U SPORTU

3. INFORMATICS IN SPORTS

Danas je moguće snimiti točne pokrete pojedinih igrača u ekipnim sportovima, kao što su nogomet, košarka, rukomet, itd.[3]

Najprodavanija knjiga Moneyball Michael Lewis promijenila je način na koji su ljudi razmišljali o sportu, posebno za vlasnike, menadžere i igrače. Lewisova knjiga pomogla je u stvaranju revolucije u kojoj se performanse igrača mjere pomoću pristupa zasnovanog na statistikama, a ne tradiciji kojom dominiraju anegdote i intuicija.

Od tada su sportski znanstvenici pokušali kopirati uspjeh ovog pristupa u sportovima poput košarke, nogometa, američkog nogometa i tako dalje. Ta je znanost potaknuta relativno novom sposobnošću prikupljanja golemih količina podataka o igračima i igri dok je igra u tijeku.

Međutim, u mnogim sportovima, sposobnost prikupljanja podataka nije podudarna s mogućnošću obrade na smislene načine. Zahvaljujući radu Joachima Gudmundssona i Michaela Hortona na Sveučilištu u Sydneyu u Australiji koji su pregledali ovo polje i istaknuli su iznimne izazove s kojima se istraživači suočavaju u analitičkom smislu.[3]

Ovaj se način prikupljanja može primijeniti na ekipne sportove u kojima se dvije momčadi sukobljavaju oko posjeda lopte na ograničenom igralištu. Svaka ekipa ima istovremeni cilj postizanja pogotka dok sprječava pogodak koji bi suparnik mogao postignuti. Dobitnik je ekipa koja postiže najveći broj pogodaka do isteka vremena.

Sportovi koji dijele tu strukturu uključuju nogomet, košarku, hokej na ledu, hokej na terenu, ragbi, rukomet, australski nogomet, američki nogomet, lacrosse itd. Međutim, većina podataka dolazi od igara poput profesionalnog nogometa i košarke koji imaju sredstva za prikupljanje tih podataka.

Ti podaci se obično sastoje od putanja igrača i lopte tijekom igre i zapisnici koji opisuju događaje kao što su dodavanja, pucanja, driblanja itd. u određenom vremenu. Danas postoji vrhunska tehnologija koja može snimati te putanje s velikom frekvencijom te se ti podaci mogu analitički obrađivati.

Veliki izazov u sportskoj znanosti jest korištenje ovih podataka kako bi se stekla konkurentna prednost, bilo u stvarnom vremenu tijekom igre, ili kako bi se pomoglo u osposobljavanju, pripremi ili regrutiranju.

Jedan od najznačajnijih problema je razumijevanje kako igrači mogu dominirati dijelovima igrališta u blizini. U sportskoj znanosti, dominantna regija igrača je regija kojoj može doći prije bilo kojeg drugog igrača. Jednostavan način za izračunavanje toga je izvući Voronoijev dijagram koji dijeli polje u regije najbliže svakom igraču. Takav dijagram može se izmijeniti pomoću drugih informacija, kao što je promatranje da dominantna područja imaju tendenciju da budu veća za napadačku momčad od obrambenog tima. Izračunavanje Voronoievog dijagrama za svakog igrača na terenu je računalno skupo jer je potrebna velika procesorska snaga kako bi se moglo računati u realnom vremenu, te time neki od timova dobio realnu prednost nad drugim. Umjesto toga, istraživači izračunavaju drugačije svojstvo što je područje koje igrač može prijeći u određenom vremenu te se preklapanja zatim rješavaju.

Čak i u tom slučaju taj sustav ignorira neke bitne čimbenike koji su inercija igrača jer igrač u pokretu može prijeći veću udaljenost i time dominirati većim područjem, nego onaj koji je stacionaran u tom istom trenutku.

To može dovesti do složenih podjela terena. Kada igrač A trči kod suprotstavljenog igrača B koji je stacionaran, svaki može imati više od jedne dominantne regije, a one možda nisu međusobno povezane. Na primjer, inercija igrača A pruža bolji pristup nekim, ali ne i svim područjima iza igrača B. Što je važan problem u sportskoj znanosti? To je kako izračunati realistične dominantne regije u realnom vremenu.

Drugi izazov s kojim se suočava je razjasniti je li igrač otvoren za dodavanje lopte. Točnije, određuje postoji li određena brzina i smjer kojom se kreće lopta kako bi igrač primio loptu prije nego što ju neki drugi igrač presretne.

To je očito povezano s dominantnom regijom igrača. S obzirom na preciznu ideju o tome što je to područje, jednostavno je riješiti pravocrtnu putanju lopte prilikom direktnog dodavanja koja se nalazi u dominantnoj regiji tog igrača, te tako trenutni alati koji izračunavaju to rade.

Problem je u tome što samo određene putanje zadovoljavaju kriterij direktnih dodavanja. Na primjer, zračna putanja nije direktno dodavanje. Još ne postoji alat koji se može nositi s tim ili s putanjama lopte koja se rotira, a to je još jedan problem u sportskoj znanosti.

Zatim postoji način na koji jedan igrač može pritisnuti druge igrače zatvarajući prostor oko sebe. Što se isto mora na određeni način prikazati.

Sve važnije područje sportske analitike uključuje mrežnu znanost. To svaki igrač tretira kao čvor i povlači se linija između njih kada lopta putuje od jednog do drugog. Ovo je plodonosno područje istraživanja jer je već razvijen širok raspon matematičkih alata za analizu mreža.

Iste vrste znanosti također omogućuju da mreža bude podijeljena u svežnjeve. Tako neki članovi tima mogu dodavati jedni drugima ili raditi učinkovitije zajedno.

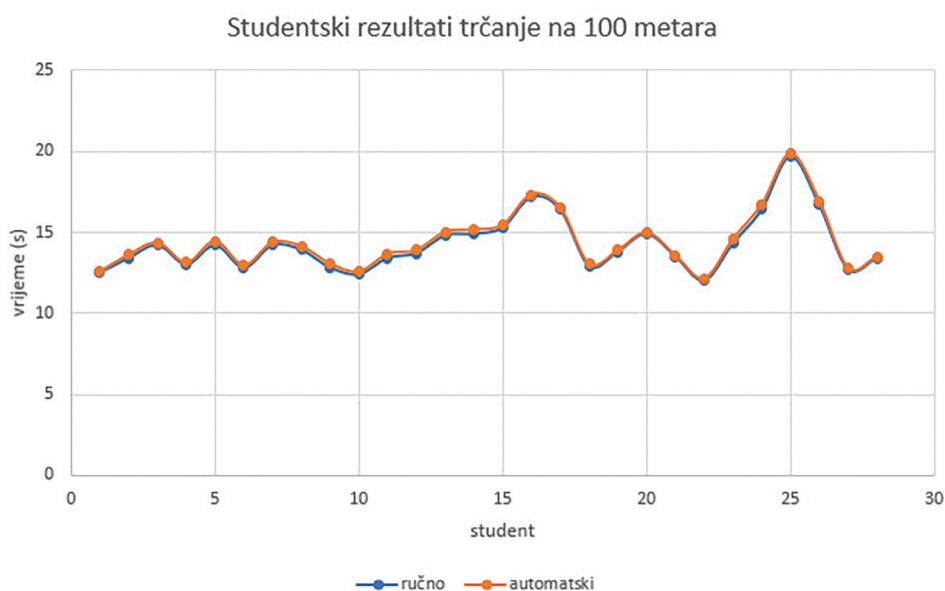
Međutim, problem sa znanostima o mreži je da postoje brojni različiti načini mjerenja centralnosti i određivanja svežnjeva, a nije uvijek jasno zašto bi jedna metoda trebala biti preferirana nad drugom. Dakle, drugi neriješeni problem je sustavno procijeniti i usporediti ove različite metode kako bi se utvrdilo njihovu korisnost i vrijednost.

Druga klasa problema dolaze iz analize podataka o igrama. Na primjer, s obzirom na popis putanja igrača i zapisnike događaja za razdoblje tijekom igre, moguće je odrediti formiranje tima - na primjer, 4-4-2 u nogometu što daje taktičku prednost.

4. MJERENJA REZULTATA U SPORTU 4. MEASUREMENTS OF RESULTS IN SPORTS

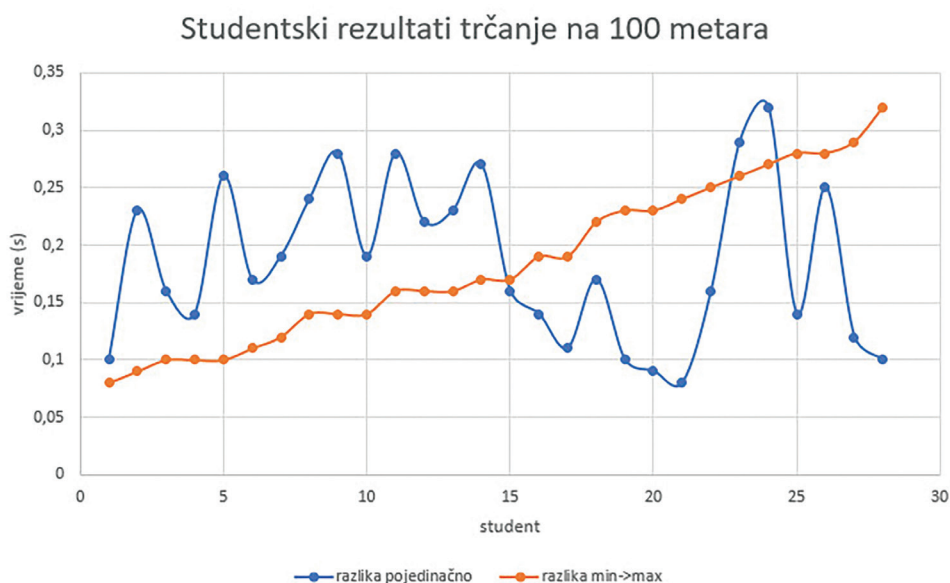
Veliko područje primjene elektronike i informatike u sportu su mjerenja postignutih rezultata i komparacija s drugim natjecateljima. Za potrebe ovog rada navest će se samo primjer razlike u mjerenjima vremena kod trkača za ručno i automatsko mjerenje vremena.

Mjerena su vremena studenata u trčanju na 100 metara. Slika 1. prikazuje dobivene rezultate ručnim mjerenjem i automatskim mjerenjem.



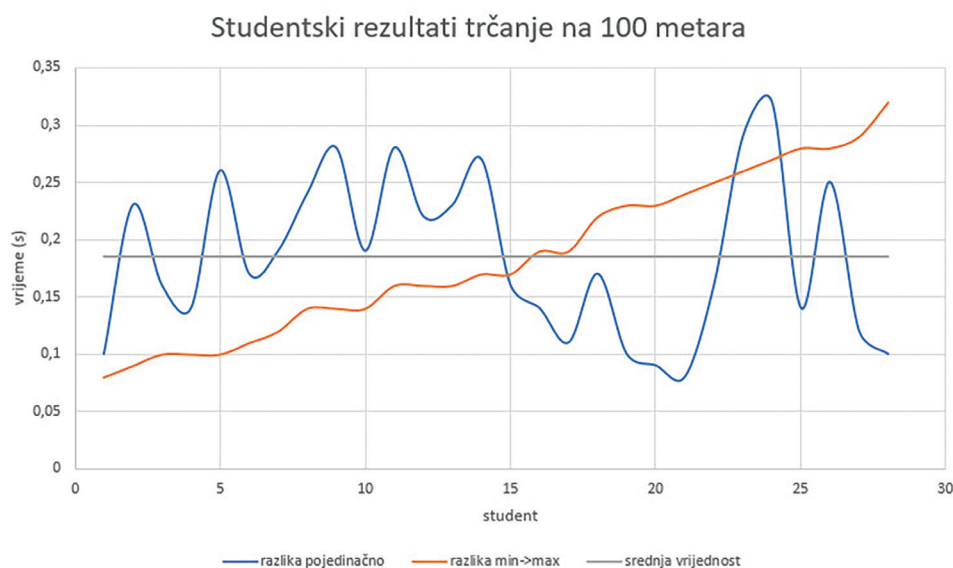
Slika 1 Rezultati na 100 metara

Figure 1 Results at 100 meters



Slika 2 Razlika u izmjerenim vrijednostima

Figure 2 Difference in measured values



Slika 3 Srednja vrijednost pogreške
Figure 3 Mean error value

Na slici 2. se može vidjeti pojedinačna razlika u greškama kod ručnog mjerenja, kao i raspon greške mjerenja od minimalne do maksimalne greške.

Prosječna pogreška u mjerenjima iznosi $0,185 \pm 0,12$ sekunde, kao što se vidi na slici 3.

Kao što se vidi iz priloženih slika, između ručnog i automatskog mjerenja nisu uvijek iste razlike iz razloga što je kod ručnog mjerenja u pitanju ljudska sposobnost i iskustvo. Kod više mjerača vremena (suci ili treneri) za istog trkača, vremena kod svakog mjerača bit će drugačija, dok je automatski rezultat uvijek isti. Naravno iskusniji suci ili treneri imat će manju razliku između ručnog i automatskog rezultata.

5. ZAKLJUČAK

5. CONCLUSION

Analitika i informatika u pojedinačnom ili ekipnom sportu su ključni alati koji mogu povećati spremnost igrača kako u ekipnim, tako i u individualnim sportovima kao i saznanje koje imaju treneri i izbornici prilikom izbora pozicija za pojedine igrače u timu, ali i za sami izbor igrača za tim koji bi najbolje funkcionirali zajedno.

Same discipline su u razvoju, ali i kao takve mogu iznimno poboljšati rad pojedine ekipe i olakšati posao stručnom stožeru u svakom koraku sportskog događaja poput sastavljanja ekipe, kako bi sastavili pojedince koji su međusobno kompatibilni na različitim područjima igre ili prilikom nabavke novog igrača kako bi nabavili najpotrebnijeg igrača. Potom prilikom priprema za pojedini susret kako bi najbolje moguće se pripremili bilo taktički bilo fizički. Te metode su iznimno korisne i prilikom samih treninga jer uz savjetovanje s istima može se postići maksimalna produktivnost pojedinog člana tima kao i produktivnost cijelog tima.

Automatska mjerenja rezultata su danas od ključne važnosti u sportu i sve više se razvijaju zahvaljujući napretku tehnologije i analitike u sportu.

6. REFERENCE

6. REFERENCES

- [1.] www.fivethirtyeight.com; 9.7.2018.;
- [2.] Silver N.; http://www.espn.com/world-cup/story/_/id/4447078/ce/us/guide-espn-spi-ratings; 2009.; 11.7.2018.;
- [3.] Gudmundsson J.; Horton M.; Spatio-Temporal Analysis of Team Sports – A Survey; 2016.;

AUTORI · *AUTHORS*

● **Valter Perinović** - nepromjenjena biografija
nalazi se u časopisu Polytechnic & Design
Vol. 2, No. 2, 2014.

Korespondencija · *Correspondence*

vperinovic@tvz.hr

● **Matej Kumiša**

Rođen je 1.12.1996. u Zagrebu, gdje završava
osnovnu školu i srednju tehničku školu (Tehnička
škola Ruđera Boškovića) smjer mehatronika.
Godine 2015. upisuje Tehničko veleučilište u
Zagrebu smjer mehatronika.

Korespondencija · *Correspondence*

matej.kumisa@tvz.hr