



Satakunnan ammattikorkeakoulu
Satakunta University of Applied Sciences

JASPER JOKINEN

Tehokas data-analytiikka: Avain NFL-vedonlyönnissä

TIETOJENKÄSITTELYN TUTKINTO-OHJELMA
2024

TIIVISTELMÄ

Jokinen, Jasper: Tehokas data-analytiikka: avain NFL-vedonlyönnissä
Opinnäytetyö, AMK
Tietojenkäsittely
Toukokuu 2024
Sivumäärä: 37

Amerikkalainen jalkapallo (NFL) kohoaa Yhdysvaltain urheiluliigojen kirkkaimmaksi tähdeksi katsojamäärissä, ja sen viihdebisnes on arvoltaan miljardeja dollareita. Otteluiden tulosten ennustaminen on suosittu aihe vedonlyöjien, akateemikkojen ja analytikoiden parissa, jotka pyrkivät ymmärtämään, mitkä tekijät vaikuttavat eniten peleihin pitkällä aikavälillä.

Tässä opinnäytetyössä tutustuttiin miten data-analytiikka voi tuoda lisäarvoa urheiluedonlyöntiin. Lisäksi käydään läpi keskeisiä käsitteitä, kuten big data -urheilu, tilastolliset mallit ja niiden soveltaminen vedonlyönnissä sekä pelaajien suoritusten analysointi. Tavoitteena on antaa syvämpi katsaus siihen, miten urheiluedonlyöjät voivat hyödyntää data-analytiikkaa omassa päätöksenteossään ja parantaa pitkän aikavälin tuloksiaan.

Opinnäytetyössä todettiin, että data-analytiikka on tehnyt urheiluedonlyönnistä tarkempaa ja kannattavampaa. Sitä voidaan käyttää monenlaisiin tarkoituksiin, kuten joukkueiden ja pelaajien suorituskyvyn analysointiin, ottelun olosuhteiden arviointiin ja arvokohteiden löytämiseen. Opinnäytetyössä myös todettiin, että urheiluedonlyönnissä data-analytiikkaa apunaan käyttäminen ei takaa voittoja.

Avainsanat: Data-analytiikka, NFL, Urheiluedonlyönti, Tilastolliset mallit, Logistinen regressio

Abstract

Jokinen, Jasper: Powerful data analytics: the key to NFL betting

Bachelor's thesis

Business Information Systems

May 2024

Number of pages: 37

National Football League (NFL) is emerging as the brightest star of US sports leagues in terms of viewership, and its entertainment business is worth billions of dollars. Predicting the outcome of matches is a popular topic among bettors, academics and analysts who seek to understand which factors will have the greatest impact on games in the long term.

This thesis explored how data analytics can add value to sports betting. It also covers key concepts such as big data sports, statistical models and their application in sports betting, and the analysis of player performance. The aim is to provide an in-depth look at how sports bettors can use data analytics in their decision-making and improve their long-term results.

The thesis found that data analytics has made sports betting more accurate and profitable. It can be used for a wide range of purposes, such as analyzing team and player performance, assessing match conditions, and finding areas of value. The thesis also found that using data analytics in sports betting does not guarantee profits.

Keywords: Data analytics, NFL, Sports betting, Statistical models, Logistic regression

SISÄLLYS

1	JOHDANTO	5
2	YLEISTIETOA AMERIKKALAISESTA JALKAPALLOSTA.....	7
2.1	Amerikkalaisessa jalkapallossa suorituksen mittaamiseen käytettävät tilastotiedot	8
3	DATA-ANALYTIIKAN TAUSTA.....	10
3.1	Big data ja urheiluedonlyönti.....	11
4	URHEILUVEDONLYÖNNIN TAUSTA	12
4.1	NFL:n vedonlyöntimarkkinoiden tila.....	13
5	DATA-ANALYTIIKAN MERKITYS URHEILUVEDONLYÖNNISSÄ.....	15
6	MALLINNUS JA ANALYSOINTI	16
6.1	Mallinnus	16
6.2	Logistinen regressio	16
6.2.1	Logistisen regression soveltaminen NFL-vedonlyönnissä.....	18
6.3	Satunnainen metsä.....	19
6.4	Tukivektorikone	20
6.5	Pelaajien suorituskyvyn analysointi	21
6.6	Kertoimien analysointi	22
6.6.1	Markkinoiden sentimentin analysointi.....	24
7	DATA-ANALYTIIKAN HAASTEET URHEILUVEDONLYÖNNISSÄ.....	25
8	PÄÄTÖKSENTEON PARANTAMINEN	27
8.1	Riskien hallinta	28
9	DATA-ANALYTIIKAN YHDISTÄMINEN VEDONLYÖNTI STRATEGIAAN 31	
10	LOPUKSI	33
	LÄHTEET.....	34

1 JOHDANTO

Urheiluviedonlyönti on suosittu ajanviettotapa ja taloudellinen harrastus monille. Päätöksenteko vedonlyönnissä on kuitenkin haastavaa, ja usein perustuu enemmän intuition kuin perusteellisen analyysin varaan. Urheiluviedonlyöntiin liittyy monia muuttujia, jotka voivat vaikuttaa tuloksiin, kuten esimerkiksi joukkueiden suorituskyky, pelaajien kunto ja sääolosuhteet.

NFL (National Football League) on Yhdysvalloissa toimiva ammattilaistasolla toimiva amerikkalaisen jalkapallon liiga. NFL on katsojamäärältään suosituin urheilulaji Yhdysvaltojen liigoista, ja se on miljardien dollarien arvoinen viihdebisnes. NFL on tunnettu intensiivisestä ja fyysisestä pelityylistään sekä laajasta kannattajakunnastaan sekä Yhdysvalloissa että ympäri maailmaa. Mahdollisuus ennustamaan otteluiden tuloksia on suosittu aihe vedonlyöjien, akateemikkojen mutta myös analyttikoiden keskuudessa, jotka haluavat selvittää, mitkä pelitapahtumat vaikuttavat eniten peleihin pitkällä aikavälillä.

Data-analytiikka on noussut keskeiseen rooliin tässä yhteydessä tarjoten urheiluviedonlyöjille mahdollisuuden hyödyntää suuria tietomassoja ja monimutkaisia algoritmeja parantaakseen vedonlyöntipäätöksiään.

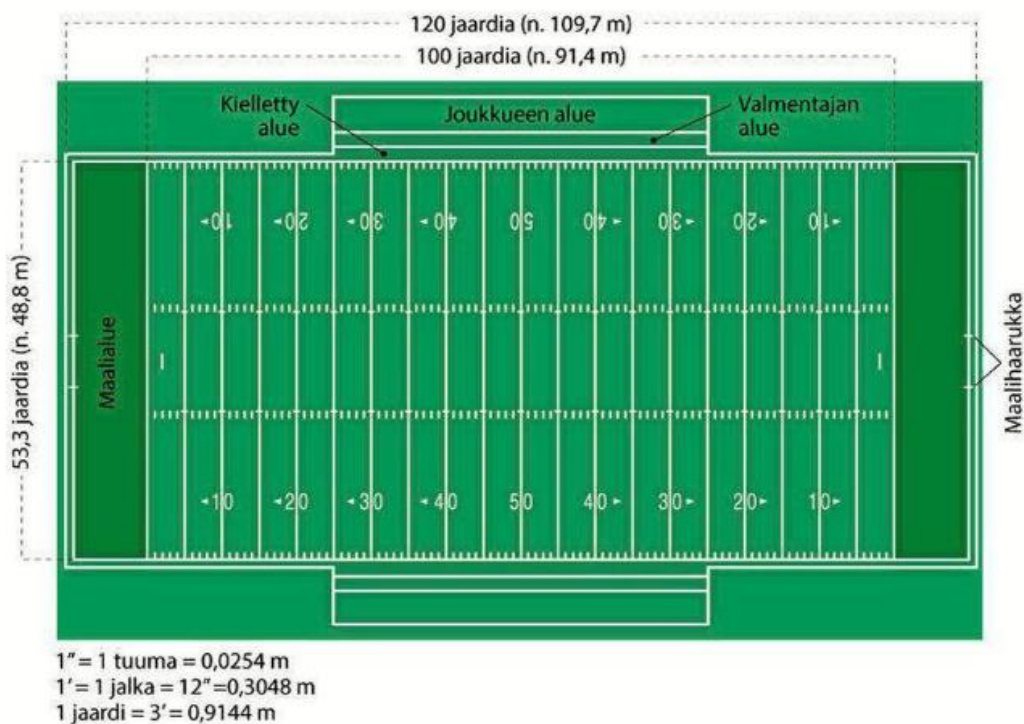
Tämän opinnäytetyön tavoitteena on tarkastella syvällisesti data-analytiikan merkitystä ja tehokkuutta NFL-viedonlyönnissä. Tarkastelen erilaisia data-analytiikan menetelmiä ja työkaluja, kuten logistista regressiota, satunnaista metsää ja tukivektorikonetta, sekä niiden soveltamista NFL-otteluiden ennustamisessa ja vedonlyöntistrategioiden kehittämisessä. Tavoitteena on antaa syvällinen katsaus siihen, miten urheiluviedonlyöjät voivat hyödyntää data-analytiikkaa omassa päätöksenteossaan ja parantaa pitkän aikavälin tuloksiaan.

Tämä opinnäytetyö tarjoaa arvokasta tietoa niille, jotka ovat kiinnostuneita hyödyntämään data-analytiikkaa NFL-vedonlyönnissä sekä alan ammattilaisille ja tutkijoille, jotka haluavat syventää ymmärrystään tästä aiheesta.

2 YLEISTIETOA AMERIKKALAISESTA JALKAPALLOSTA

Amerikkalainen jalkapallo on joukkueurheilulaji, jota pelataan kahden joukkueen välillä. Jokaisella joukkueella on tavoitteena saada enemmän pisteitä kuin vastustajalla. Peli pelataan neljännesvuoroissa, joissa yritetään edetä kentällä ja tehdä maaleja. Peli alkaa potkulla, ja joukkueet vuorottelevat hyökkäyksen ja puolustuksen välillä. Tärkeimpiä pelaajia ovat pelinrakentaja, jonka tehtävänä on ohjata hyökkäystä, sekä linjamiehet, jotka suojaavat häntä. Pisteitä voi saada esimerkiksi tekemällä maalin (touchdown), potkaisemalla pallon maalipotkulla (field goal) tai saamalla vastustajan rangaistusalueella pelinrakentajan pysähtymään (safety). Peli päättyy yleensä neljännesvuoron tai jatkoajan jälkeen, ja voittajajoukkue on se, jolla on enemmän pisteitä.



Amerikkalaisen jalkapallon kenttä



Kuva 1. Amerikkalaisen jalkapallon kenttä. (Degerlund, 2022)

2.1 Amerikkalaisessa jalkapallossa suorituksen mittaamiseen käytettävät tilastotiedot

Amerikkalaisen jalkapallon suorituskyvyn mittaamiseksi käytetään monia erilaisia tilastoja ja mittareita. Tärkeimpiä näistä ovat yardit, jotka kertovat pelaajien tai joukkueiden etenemästä, touchdownit, jotka ilmaisevat joukkueen kyvyn päästä vastustajan maalialueelle, interceptionit ja syötönkatkot, jotka mittaavat puolustuksen suorituskykyä, säkitykset, jotka osoittavat puolustuksen kyvyn estää vastustajan hyökkäys, kolmannen alueen onnistumisprosentti, joka mittaa onnistuneiden pelisuunnitelmien määrää, sekä pelikellon hallinta, joka osoittaa joukkueen kyvyn hallita pelin kulkua ja aikaa. Näitä tilastoja hyödynnetään valmentajien, pelaajien ja fanien toimesta pelin dynamiikan ja joukkueiden suorituskyvyn ymmärtämiseksi. Alla on esimerkki kuva San Francisco 49ers ja Kansas City Chiefs välisestä ottelusta kerätyistä tilastoista.

Team Stats		
		
1st Downs	23	24
Passing 1st downs	15	15
Rushing 1st downs	5	9
1st downs from penalties	3	0
3rd down efficiency	3-12	9-19
4th down efficiency	1-1	1-1
Total Plays	71	79
Total Yards	382	455
Total Drives	13	13
Yards per Play	5.4	5.8
Passing	272	325
Comp-Att	24-39	34-46
Yards per pass	6.8	6.6
Interceptions thrown	0	1
Sacks-Yards Lost	1-4	3-8
Rushing	110	130
Rushing Attempts	31	30
Yards per rush	3.5	4.3
Red Zone (Made-Att)	1-2	2-6
Penalties	6-40	6-55
Turnovers	2	2
Fumbles lost	2	1
Interceptions thrown	0	1
Defensive / Special Teams TDs	0	0
Possession	38:31	36:26

Kuva 2. Super Bowl LVIII:n tilastot (ESPN, 2024)

3 DATA-ANALYTIIKAN TAUSTA

2000-luvun alun myötä käsite data-analytiikka nousi suosioon. Data-analytiikassa keskitytään suurten tietomassojen, eli niin kutsutun big datan, analysointiin ja päätöksenteon tukemiseen tietokonejärjestelmien avulla. Tämä ala on hyvin monitieteinen, ja siihen on integroitu näkökulmia useilta eri tieteenaloilta, kuten tilastotiede, koneoppiminen, mallintaminen, järjestelmäteoriatiede ja tekoäly. (Runkler, 2016, s.26)

Data-analytiikka on prosessi, jossa käsitellään ja tulkitaan tietoa (dataa) sen hyödyllisen tiedon ja oivallusten saamiseksi. Se on laaja käsite, joka kattaa monenlaisia toimintoja ja menetelmiä, ja sitä sovelletaan monilla eri aloilla, kuten liike-elämässä, tieteessä, terveydenhuollossa ja teknologiassa.

The Investopedia Team (2024) listaa artikkelissaan "Data Analytics: What It Is, How It's Used, and 4 Basic Techniques" laajan katsauksen data-analytiikan keskeisiin osa-alueisiin:

- **Datankeruu:** Ensimmäinen vaihe data-analytiikassa on datan kerääminen. Tämä voi tapahtua monilla eri tavoilla, kuten manuaalisesti syöttämällä tietoja järjestelmiin, automaattisesti keräämällä tietoja antureista tai IoT-laitteista, tai keräämällä tietoja verkosta tai sosiaalisesta mediasta.
- **Datan puhdistus ja esikäsittely:** Usein raakadata sisältää virheitä, puutteita tai epätarkkuuksia, joten se on puhdistettava ja esikäsiteltävä ennen kuin sitä voidaan käyttää analytiikassa. Tähän kuuluu esimerkiksi virheellisten tietojen poistaminen, puuttuvien tietojen täydentäminen ja datan muuntaminen sopivaan muotoon analysointia varten.

- Tietojen analysointi: Kun data on puhdistettu ja esikäsitelty, sitä voidaan analysoida erilaisilla menetelmillä ja työkaluilla. Tähän kuuluvat tilastolliset analyysit, koneoppiminen, tekoäly ja data mining -tekniikat. Tavoitteena on löytää merkityksellisiä kuvioita, suuntauksia ja yhteyksiä datasta, jotka voivat auttaa päätöksenteossa ja ennustamisessa.
- Tulkinta ja raportointi: Analyysin tulosten tulkinta on tärkeä osa data-analytiikkaa. Tuloksista tehdään raportteja, visualisointeja ja kuvaajia, jotka auttavat kommunikoimaan löydökset sidosryhmille ymmärrettävästi ja vaikuttavasti. Tulkinnassa pyritään ymmärtämään, mitä data kertoo ja mitä toimenpiteitä tulisi toteuttaa sen perusteella.

Data-analytiikka on prosessi, joka hyödyntää dataa päätöksenteossa, ongelmien ratkaisemisessa ja tulevaisuuden ennustamisessa. Se on olennainen osa modernia liiketoimintaa ja yhteiskuntaa, ja sen merkitys kasvaa jatkuvasti digitalisaation ja teknologian kehityksen myötä.

3.1 Big data ja urheiluvedonlyönti

Vielä muutamia vuosia sitten oli vaikea kuvitella, että kykenisimme keräämään ja analysoimaan yhtä mittavia tietomääriä kuin nykyään. Tänä päivänä käytössämme oleva tietomäärä on valtava, ja vuoteen 2020 mennessä ennustettiin, että joka sekunti syntyy 1,7 megatavua uutta tietoa. Suurien data-massojen keräämistä ja analysointia kutsutaan termillä big data. (Bernard, 2015, s.1-2).

Big data–analytiikka mahdollistaa massiivisen määrän datan tehokkaan käsittelyn. Tämä antaa vedonlyöjille mahdollisuuden tunnistaa piileviä trendejä. Itotian (2023) mukaan esimerkiksi pelaajien fyysisten ominaisuuksien ja suoritus-
ten syvälinen analyysi voi paljastaa yhteyksiä ja auttaa tekemään parempia vedonlyöntipäätöksiä

4 URHEILUVEDONLYÖNNIN TAUSTA

Nykyään urheiluedonlyönti on laajalle levinnyt ilmiö, joka kattaa laajan valikoiman urheilulajeja ja tapahtumia ympäri maailmaa. Se on kasvanut merkittävästi viime vuosikymmeninä, osittain teknologisten edistysaskelten ja vedonlyöntiteollisuuden laajentumisen ansiosta. Vedonlyöntimarkkinat ovat monimutkaistuneet ja tarjoavat monia erilaisia palveluita ja vaihtoehtoja vedonlyöjille. (Pempus & Plantana, 2024)

Urheiluedonlyönti on nykyään miljardien dollarien teollisuus. Vedonlyöntiyrietykset toimivat kaikkialla maailmassa, ja ne tarjoavat laajan valikoiman tapahtumia ja kertoimia. Urheiluedonlyönnistä on tullut suosittu harrastus, ja sitä nauttivat ihmiset kaikenikäisistä ja -taustaisista. Yksi merkittävimmistä kehityskuluista on ammattimaisen vedonlyönnin nousu, jossa ihmiset elättävät itsensä vedonlyönnillä kokopäiväisesti. Ammattimaiset vedonlyöjät käyttävät edistyneitä tilastollisia ja matemaattisia malleja, koneoppimista ja muita analyysityökaluja ennustaakseen otteluiden tuloksia ja löytääkseen aliarvostettuja kertoimia.

Urheiluedonlyönnissä osallistuja asettaa valitsemansa panoksen vedonvälittäjän tarjoamaan kohteeseen. Kohteen lopputulos ratkeaa usein saman päivän aikana. Vedonvälittäjät määrittelevät jokaiselle tapahtumalle kertoimet, jotka heijastavat vedonlyöntisivuston arviota tapahtuman todennäköisyydestä. Pienempi kerroin ilmaisee vedonvälittäjän näkemyksen siitä, että kyseinen tapahtuma on todennäköisempi toteutumaan.

Urheiluedonlyönti perustuu usein intuition ja tietämykseen, mutta pelkkä tieto ei välttämättä ole riittävästi. Data-analytiikka voi tarjota syvällisempää ymmärrystä joukkueiden suorituskyvystä, pelaajien tilastoista ja pelien dynamiikasta. Tämä tieto auttaa vedonlyöjää tekemään perustellumpia päätöksiä. (Itotia, 2023).

4.1 NFL:n vedonlyöntimarkkinoiden tila

Vedonlyöntimarkkinat ovat mielenkiintoinen tutkimusaihe, koska taloudellisen hyödyn motivaatio todennäköisesti kannustaa luomaan mahdollisimman tarkkoja ennusteita ja asiantuntemusta, ja voidaan väittää, että tämä saisi vedonlyöntimarkkinoita tehokkaaksi. Tutkijat ovat tarkastelleet NFL:ää eri näkökulmista nähdäkseen, onko näin.

Borghesi (2007) tarkasteli sääolosuhteiden vaikutusta peleihin. Sää ja lämpötila vaikuttavat merkittävästi amerikkalaisessa jalkapallossa, varsinkin kuinka helppoa on heittää ja ottaa kiinni palloa. Tutkimusartikkelissa Borghesi tarkasteli, voidaanko jonkin joukkueen katsoa hyötyvän säästä. Tutkimuksessa havaittiin, että äärimmäisissä lämpötiloissa pelattaviin peleihin liittyy väärin hinnoiteltu sopeutumisetu, ja vedot kylmään sopeutuneesta kotijoukkueesta, joka pelaa omassa elementissään, ovat huomattavasti alihinnoiteltuja. Nämä havainnot viittaavat siihen, että NFL-vedonlyöntimarkkinat eivät ole täysin tehokkaat ja, että siellä voi olla mahdollisuuksia kannattaviin vedonlyöntistrategioihin, jotka perustuvat säähän liittyviin harhoihin.

Samainen kirjoittaja otti erilaisen näkökulman toisessa tutkimuksessaan, missä hän tutki ilmiötä, jossa kotijoukkue on altavastaajana NFL:n vedonlyöntimarkkinoilla. Hän havaitsi, että tämän ilmiön ensisijainen syy oli ennakoasenteiden voimakas kasvu kunkin kauden viimeisillä viikoilla. Kirjoittaja hyödynsi useita tilastollisia testejä ja regressioanalyysyjä vahvistaakseen ilmiön olemassaolon ja selvittääkseen sen syitä. Hän osoitti, että NFL-vedonlyöntikertoimet poikkeavat jatkuvasti niiden todellisista arvoista tietyn 20 vuoden ajanjakson aikana. Lisäksi kirjoittaja esitti useita regressiomalleja, joita voidaan käyttää tehottomuuksien tunnistamiseen ja virheellisen hinnoittelun ennakoimiseen ennen niiden tapahtumista. Malli eroaa aiemmista kirjallisuudessa käytetyistä malleista, koska se ottaa huomioon sekä lyhyen aikavälin että kokonaishäiriöt. Yhteenvetona voidaan todeta, että tutkimus tarjoaa

uuden lähestymistavan virheellisen hinnoittelun tunnistamiseen NFL:n vedonlyöntimarkkinoilla sekä voittoa tavoittelevien vedonlyöntijärjestelmien toteuttamiseen. (Borghesi, 2011)

5 DATA-ANALYTIIKAN MERKITYS URHEILUVEDONLYÖNNISSÄ

Data-analytiikka on muuttanut urheiluedonlyöntiä monin tavoin, kuten teke-mällä vedonlyönnistä tarkempaa. Data-analytiikkaa käyttämällä vedonlyöjät voivat analysoida suuria määriä tietoa. Näin he voivat myös tehdä parempia päätöksiä siitä, mihin vedonlyöntiä sijoittavat. (Jaiswal, 2023).

Data-analytiikkaa voidaan käyttää joukkueiden ja pelaajien suorituskyvyn analysointiin. Tämä voidaan tehdä analysoimalla esimerkiksi joukkueiden voittoja ja tappioita, maalimääriä, pelaajien pisteitä ja muita tilastoja. (Harrison & Bukstein, 2016)

NFL urheilutapahtumissa esiintyy usein kausiluonteisia vaihteluita, kuten sää-olosuhteiden vaikutusta ulkourheilulajeissa tai joukkueiden suorituskyvyn vaihteluita kauden eri vaiheissa. Data-analytiikan avulla voidaan tunnistaa nämä kausiluonteiset vaihtelut ja ottaa ne huomioon vedonlyöntistrategioissa ja ennusteissa. Analysoimalla suurta määrää historiallista dataa urheilutapahtumista, voidaan havaita toistuvia kuvioita ja trendejä, jotka auttavat vedonlyöjiä ymmärtämään paremmin urheilumarkkinoiden dynamiikkaa ja teke-mään informoituja päätöksiä vedonlyöntimarkkinoilla. (Solar, 2015)

6 MALLINNUS JA ANALYSOINTI

6.1 Mallinnus

Tilastolliset mallit ovat keskeisessä roolissa, kun pyritään ennustamaan urheilutapahtumien tuloksia Altmannin (2004, s.18) mukaan tilastolliset mallit voivat hyödyntää pelaajientilastoja, joukkueiden suorituskykyä ja muita peliin liittyviä tekijöitä ennustaakseen tulevia tuloksia. Tällainen lähestymistapa voi tarjota objektiivista tietoa, joka ylittää ihmisen intuition.

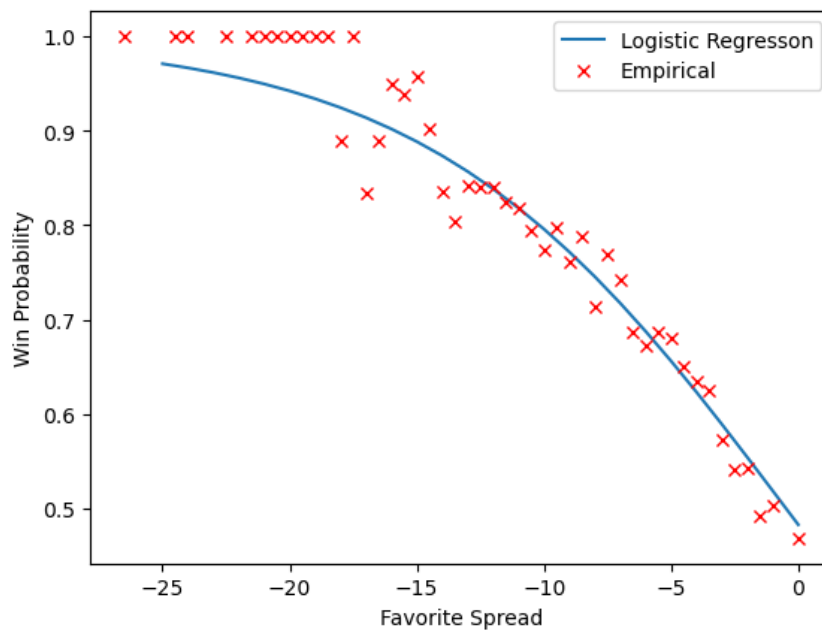
Datapohjaisessa urheiluedonlyönnissä keskiössä ovat tilastolliset mallit, joiden laskenta perustuu historiatietoihin. Nämä mallit arvioivat tietyn pelin tulosten todennäköisyydet. Kvantifioimalla nämä todennäköisyydet, voimme vertailla niitä vedonvälittäjien kertoimiin. Tällä tavoin pystymme tunnistamaan mahdolliset erot ja löytämään arvokohteen, joka voi tarjota pelaajalle mahdollisuuden hyötyä vedonlyönnissä. (Jaiswal, 2023). Arvokohde on veto, jonka todennäköisyys on suurempi kuin sen vedonlyöntikerroin. Datan avulla vedonlyöjät voivat tunnistaa vetoja, joissa on suuri todennäköisyys voittoon, mutta alhainen vedonlyöntikerroin. (Bookkerit, n.d.)

Tilastollinen analyysi on jatkuva prosessi. Kun uutta tietoa kerätään, tilastollisia malleja on mukautettava ja jalostettava. Säännöllinen seuranta ja mallin säädöt voivat tukea ennakoitokykyäsi dynaamisella urheiluedonlyönnillä. (Walter, 2023)

6.2 Logistinen regressio

Logistinen regressio on tilastollinen malli, jota käytetään ennustamiseen. Se soveltuu luokitusongelmiin, sillä lineaarinen regressio soveltuu paremmin jatkuvien arvojen, kuten iän, rahamäärän ja muiden vastaavien arvojen ennustamiseen. (Nyblom, 2015, s. 38-40)

Logistista regressiota on sovellettu urheiluedonlyönnissä Silvermanin ja Suchardin (2012) toimesta, jotka hyödynsivät ehdollista logistista regressio-menetelmää Hongkongin hevoskilpailujen voittajan ennustamiseen. Tutkimuksessa käytettiin 3681 hevoskilpailun historiallisia tietoja, joista jokaisesta hevosesta kerättiin 186 tilastollista muuttujaa, kuten lepopäivät, viimeisin kilpailu, ratsastajan painonmuutos, aiempien kilpailujen keskinopeus ja se, oliko hevonen muuttanut painoaan viime kilpailun jälkeen. Mallin kouluttamiseen käytettiin 2994 kilpailua, kun taas loput 737 kilpailua (20 % aineistosta) käytettiin mallin testaamiseen. Silverman & Suchard (2012) saavuttivat mallin avulla urheiluedonlyönnissä 36,73 %:n enimmäistuoton, mikä osoittaa mallin olevan varsin menestyksekkäs tähän tiettyyn käyttötarkoitukseen sovellettuna.



Kuva 3. Logistisen regressio mallin esimerkki urheiluedonlyönnissä (Jdashbrock, n.d)

6.2.1 Logistisen regression soveltaminen NFL-vedonlyönnissä

Tiettyjen tilastollisten menetelmien, kuten logistisen regression, soveltaminen NFL-vedonlyönnissä voi auttaa ennustamaan otteluiden tuloksia. Logistinen regressio on tilastollinen menetelmä, jota käytetään selittävänä analyysimenetelmänä, kun selitettävä muuttuja on binääristä luokkaa. Tämä tarkoittaa, että se soveltuu hyvin esimerkiksi voitto-tappio-tilanteiden ennustamiseen, kuten urheiluotteluissa tapahtuvaan vedonlyöntiin.

Peoples (2024) kertoo artikkelissaan ”Playbook Projections: Implementing Logistic Regression to Predict NFL Game Outcomes”, miten logistista regressiota voidaan soveltaa NFL-vedonlyöntiin:

- **Datakeruu:** Kerää dataa NFL-otteluiden historiallisista tuloksista sekä erilaisista tilastoista, kuten joukkueiden ja pelaajien suorituskyvystä, loukkaantumisista, sääolosuhteista jne.
- **Muuttujien valinta:** Valitse muuttujat, joita haluat käyttää logistisen regressiomallin selittävinä muuttujina. Näitä voivat olla esimerkiksi joukkueiden voittoprosentti, piste-erojen keskiarvot, hyökkäys- ja puolustuskeskiarvot, pelaajien loukkaantumiset, koti- ja vierasotteluiden tilastot jne.
- **Datan esikäsittely:** Käsittele dataa tarvittaessa, kuten puuttuvien arvojen täyttäminen, skaalaaminen tai muu normalisointi.
- **Logistisen regressiomallin sovittaminen:** Sovita logistinen regressiomalli valituilla selittäville muuttujilla ja selitettävällä muuttujalla, esimerkiksi ottelun lopputuloksella (voitto/tappio).
- **Mallin testaus ja arviointi:** Testaa mallin suorituskykyä käyttämällä sitä ennustamaan uusien otteluiden tuloksia. Arvioi mallin suorituskykyä

käyttämällä erilaisia arviointimenetelmiä, kuten tarkkuutta (accuracy), F1-pistettä jne.

- Optimointi ja parantaminen: Tarvittaessa hienosäädä mallia ja kokeile erilaisia muuttujia ja ominaisuuksia parantaaksesi ennustusten tarkkuutta.
- Tulosten tulkinta: Tulkkaa mallin tuloksia ja selvittää, mitkä tekijät vaikuttavat eniten otteluiden tuloksiin. Tämä voi auttaa ymmärtämään, mitkä tekijät ovat tärkeimpiä menestyksekkään vedonlyönnin kannalta.

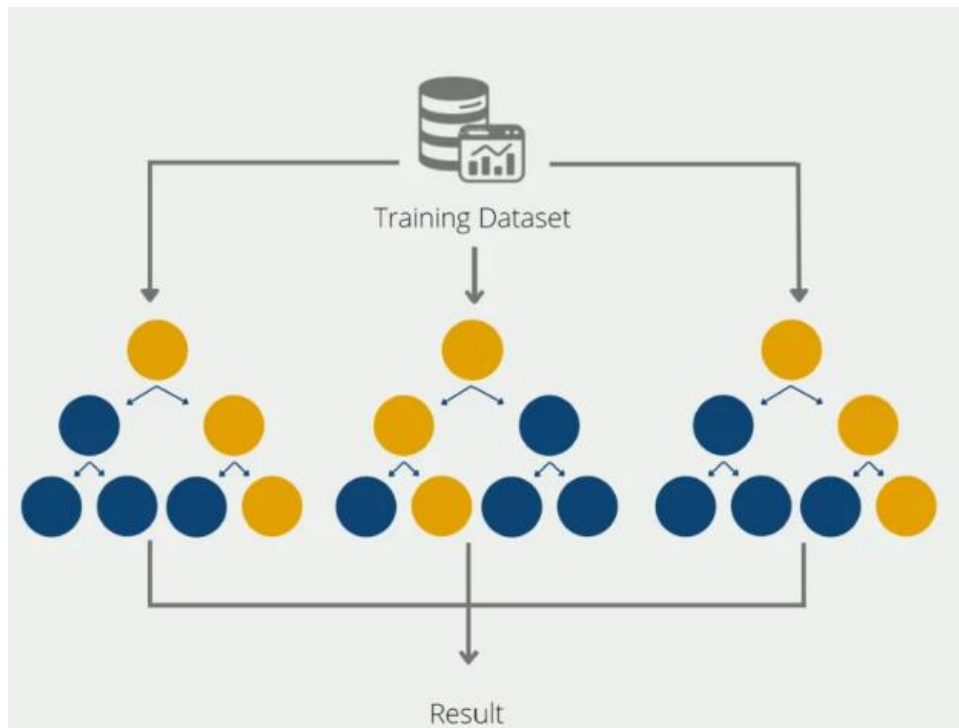
On kuitenkin tärkeää muistaa, että logistinen regressio on yksi monista mahdollisista ennustamismenetelmistä, ja sen suorituskyky riippuu suuresti käytetystä datasta ja valituista selittävästä muuttujista.

6.3 Satunnainen metsä

Satunnainen metsä toimii kuin joukko yksittäisiä päätöspuita, jotka yhdessä muodostavat kokonaisen metsän. Jokainen yksittäinen puu käyttää osaa kokonaisesta koulutusaineistostaan ja satunnaisesti valitsee piirteen, jonka perusteella se jakaa datan eri ryhmiin. Näiden puiden tulokset yhdistetään sitten enemmistöäänestyksellä, jolloin saadaan satunnaisen metsän lopputulos. Tällä tavalla satunnainen metsä hyödyntää monipuolisesti erilaisia tietoja ja piirteitä sekä vähentää yksittäisten puiden taipumusta ylisovittamiseen. (Shalev-Shwartz & Ben-David 2014, s.250-258)

Satunnaiset metsät voivat auttaa ennustamaan urheilutapahtumien tuloksia, kuten otteluiden voittajia tai pistemääriä. Näitä malleja koulutetaan historiallisilla urheiludatoilla, mukaan lukien joukkueiden suorituskyky, pelaajien tilastot, pelipaikan olosuhteet jne. Satunnaiset metsät tarjoavat luotettavia ennusteita ja voivat parantaa vedonlyöjän mahdollisuuksia menestykseen. (Bakirov, 2021). Satunnaiset metsät tarjoavat tehokkaan ja monipuolisen työkalun

urheiluviedonlyöntiin, joka voi auttaa vedonlyöjiä tekemään parempia päätöksiä ja optimoimaan panostuksensa.



Kuva 4. Satunnaisen metsän rakenne (Databasecamp, 2022)

6.4 Tukivektori-kone

Tukivektori-koneet (Support Vector Machines eli SVM) ovat koneoppimismenetelmä, jota käytetään luokitteluongelmien ratkaisemiseen. Ne pyrkivät löytämään optimaalisen päätösratkaisun eri luokkien välillä. SVM-menetelmä perustuu erottavan hyperpinnan etsimiseen, joka jakaa datan luokkiin mahdollisimman selvästi. Tämä hyperpinta on sellainen, että se maksimoi luokkien välisen marginaalin, eli etäisyyden lähimmän datapisteen ja hyperpinnan välillä. (Nelson, 2020)

SVM:tä voidaan soveltaa monenlaisiin luokitteluongelmiin, ja urheiluviedonlyöntissä, kuten NFL-viedonlyöntissä, sitä voidaan käyttää esimerkiksi pelin lopputuloksen ennustamiseen tai joukkueiden suorituskyvyn arviointiin. Tähän käytetään usein historiallista tilastodataa, kuten pelaajien tilastoja,

joukkueiden suorituskyyä eri olosuhteissa ja muita relevantteja tekijöitä. (Krueger, 2021a)

SVM:iä voidaan kouluttaa historiallisten tietojen avulla ennustamaan tulevien pelien tuloksia. Tämä voi auttaa vedonlyöjiä tunnistamaan aliarvostettuja joukkueita ja tekemään kannattavia vetoja. (Krueger, 2021b)

6.5 Pelaajien suorituskyyän analysointi

Data-analytiikkaa voidaan käyttää arvioimaan pelaajien suorituskyyä eri tilanteissa, kuten koti- ja vierasotteluissa, eri sääolosuhteissa tai eri vastustajia vastaan. Vedonlyöjät voivat analysoida pelaajien tilastotietoja ja suorituskyyä eri konteksteissa saadakseen paremman käsityksen siitä, miten he todennäköisesti suoriutuvat tulevissa otteluissa. (Soni, 2023)

Bakirov ym. (2022) suorittivat tutkimuksen amerikkalaisen jalkapallon loukkaantumisten ennustamisesta koneoppimisen avulla. Tutkijat totesivat, että pyrkimykset ymmärtää harjoittelun ja kilpailun kuormituksen sekä loukkaantumisten välistä suhdetta ovat ratkaisevan tärkeitä jalkapallossa, jotta voidaan auttaa ymmärtämään harjoitusohjelmiin sopeutumista, arvioida väsymystä ja palautumista sekä minimoida loukkaantumis- ja sairastumisriskiä. Teknologiset edistysaskeleet ovat mahdollistaneet monipisteisten tietojen keräämisen analyysejä ja loukkaantumisten ennustamista varten.

Tutkimuksessa havaittiin, että vaikka koneoppimisella on potentiaalia tuoda uutta ymmärrystä amerikkalaisen jalkapallon loukkaantumisten ennustamiseen, sen käytössä on useita rajoituksia ja haasteita. Yksi merkittävä huolenaihe on, että tutkitut tutkimukset perustuvat yhden kauden aikana kerättyihin tietoihin, mikä voi rajoittaa niiden yleistettävyyttä muihin kausiin tai konteksteihin. Lisäksi tutkimussuunnittelussa ja analyysissä on huomattavaa vaihtelua, mikä voi vaikuttaa vammojen ennustemallien tarkkuuteen. Koneoppimisalgoritmit tarvitsevat tehokkuutensa vuoksi suuria määriä korkealaatuisia tietoja, mikä voi olla käytännössä haastavaa saada.

6.6 Kertoimien analysointi

Data-analytiikka tarjoaa vedonlyöjille tehokkaita työkaluja kertoimien analysointiin ja markkinoiden trendien tunnistamiseen. Tilastollisia malleja ja historiallisia tietoja hyödyntämällä vedonlyöjät voivat arvioida, ovatko tietyt kertoimet oikeudenmukaiset vai vinoutuneet. (Odds Matrix, 2023)

Kertoimien analysointi on tärkeä osa vedonlyöntistrategiaa, erityisesti kun kyseessä on NFL-vedonlyönti. Vedonlyöntikertoimet ovat markkinoiden odotuksia ottelun lopputuloksesta ja niiden perusteellinen analysointi voi auttaa tunnistamaan arvoa, eli tilanteita, joissa kertoimet eivät heijasta todennäköisyyksiä oikein. Alla on esimerkki, miten vedonlyönti kertoimia voidaan analysoida.

Monte Carlo -simulaatio on todennäköisyyksimenetelmä, jota käytetään monimutkaisten järjestelmien mallintamiseen ja niistä saatavien satunnaisten tulosten ennustamiseen. Se toimii toistamalla satunnaiskokeita useita kertoja ja laskemalla jokaisen kokeen tulokset. Tämän avulla voidaan arvioida todennäköisyyttä eri tapahtumille tai tuloksille. (Iordanova, 2022)

Monte Carlo -simulaatiot ovat tehokas työkalu monimutkaisten järjestelmien mallintamiseen ja niistä saatavien satunnaisten tulosten ennustamiseen. Niitä käytetään monilla eri aloilla ja niillä on useita etuja, kuten monipuolisuus, helppokäyttöisyys ja tehokkuus. On kuitenkin tärkeää olla tietoinen niiden rajoituksista, kuten laskennallinen kalleus ja riippuvuus satunnaisista syötteistä. (Kenton, 2023)

```
import random

# Vedonlyöntikertoimet
kotivoitto = 2.0
tasapeli = 3.5
```

```

vierasvoitto = 3.0

# Simulaatioiden lukumäärä
simulaatiot = 10000

# Voittojen laskeminen
voitot = 0

for _ in range(simulaatiot):
    # Tuloksen generointi
    tulos = random.random()

    # Voiton laskeminen
    if tulos < 1/kotivoitto:
        voitot += kotivoitto - 1
    elif tulos < (1/kotivoitto) + (1/tasapeli):
        voitot += tasapeli - 1
    else:
        voitot += vierasvoitto - 1

# Todennäköisen voiton laskeminen
todennakoinen_voitto = voitot / simulaatiot

# Tulosta tulokset
print(f"Todennäköinen voitto: {todennakoinen_voitto}")

```

Ohjelma 1. Vedonlyöntikertoimien analysointi Monte Carlo –simulaatiolla

Koodi alkaa määrittelemällä vedonlyöntikertoimet kotivoitolle, tasapelille ja vierasvoitolle. Sitten se määrittää, kuinka monta kertaa simulointi suoritetaan (simulaatiot). Itse simulaatio tapahtuu for-silmukan avulla, joka suoritetaan simulaatioiden määrän verran. Joka kierroksella generoidaan satunnaisluku (tulos) välillä 0 ja 1 käyttäen random-moduulia.

Tämän jälkeen koodi arvioi, mikä lopputulos ottelusta on tällä satunnaisluvulla, ja määrittää sen perusteella voiton. Esimerkiksi, jos generoitu luku on pienempi kuin 1/kotivoitto, se tarkoittaa kotivoittoa, ja voittojen summaan lisätään kotivoiton kerroin miinus yksi (kotivoitto - 1). Samoin tehdään tasapelin ja vierasvoiton kohdalla.

Kun kaikki simulaatiot on suoritettu, lasketaan todennäköinen voitto jakamalla voittojen summa simulaatioiden määrällä. Tämä antaa arvion keskimääräisestä voitosta yhdellä vedonlyönnillä. Lopuksi tulostetaan todennäköinen voitto. Käytännössä tämä koodi auttaa arvioimaan, kuinka paljon voittoa keskimäärin odotetaan tietyn määrän vedonlyöntien perusteella, olettaen että vedot tehdään satunnaisesti ja että vedonlyöntikertoimet ovat oikeudenmukaiset.

6.6.1 Markkinoiden sentimentin analysointi

Markkinoiden sentimentti viittaa markkinoiden yleiseen mielipiteeseen tietyn osakkeen, sijoitusvälineen tai markkinan arvosta. Se voi olla positiivinen, negatiivinen tai neutraali. Sentimentin analysointi voi tarjota arvokasta tietoa NFL-vedonlyönnissä. Seuraamalla sosiaalisen median keskusteluja ja online-foorumeita voi saada käsityksen yleisestä mielipiteestä tietyistä joukkueista, pelaajista tai peleistä. Analysoimalla tätä tietoa ja vertaamalla sitä vedonlyöntikertoimiin voi tunnistaa tilanteita, joissa yleinen mielipide poikkeaa kertoimista, mikä voi tarjota arvokkaita vedonlyöntimahdollisuuksia. (Brown ym, 2017)

Sosiaalisen median alustoilla, kuten Twitterissä, Facebookissa ja Redditissä, käydään laajaa keskustelua sijoituksista, yrityksistä ja taloudesta yleisesti. Data-analytiikkaa voidaan käyttää analysoimaan tätä keskustelua ja tunnistamaan trendejä, mielipiteitä ja tunnetiloja.

7 DATA-ANALYTIIKAN HAASTEET URHEILUVEDONLYÖNNISSÄ

Vaikka urheiluedonlyönnissä data-analytiikan mahdollisuudet ovat valtavat, on olemassa tiettyjä kysymyksiä ja rajoituksia, jotka on otettava huomioon. Tietojen laatu, saavutettavuus ja tulkinta ovat kriittisiä näkökohtia, jotka voivat vaikuttaa ennusteiden tarkkuuteen. Lisäksi urheilutapahtumien luontainen arvaamattomuus tarkoittaa, että monimutkaisimmatkin mallit eivät ole täysin tarkkoja. (Walter, 2023)

News Flash (2024) listaa artikkelissa: "The Impact of Data Analytics on Sports Betting Strategies" data-analytiikan mahdollisia haasteita urheiluedonlyönnissä.

- Tiedon laatu ja saatavuus

Yksi keskeisimmistä haasteista urheiluedonlyönnissä on tiedon laatu ja saatavuus. Vaikka erilaisia tilastoja ja suoritusdataa on saatavilla, osa siitä voi olla epätarkkaa tai puutteellista. Lisäksi joissain urheilulajeissa, kuten NFL:ssä, tietoa voi olla vaikea kerätä ja analysoitavaa dataa saattaa olla rajoitetusti saatavilla.

- Monimutkaisuus ja vuorovaikutukset

Urheiluedonlyöntiin vaikuttavat monet eri tekijät, ja näiden tekijöiden väliset vuorovaikutukset voivat olla hyvin monimutkaisia. Esimerkiksi yksittäisen pelaajan loukkaantuminen voi vaikuttaa koko joukkueen suorituskykyyn, ja samalla sääolosuhteet voivat vaikuttaa pelin kulkuun ja lopputulokseen. Näiden monimutkaisten vuorovaikutusten mallintaminen ja ennustaminen voi olla haastavaa data-analytiikan näkökulmasta.

- Yllätyksellisyys ja epävarmuus

Urheiluotteluissa tapahtuu usein odottamattomia käänteitä ja yllätyksiä, mikä lisää epävarmuutta ennustamisessa. Vaikka data-analytiikka voi auttaa tunnistamaan trendejä ja kaavoja historiallisesta datasta, se ei voi täysin ennustaa odottamattomia tapahtumia tai poikkeuksellisia tilanteita, jotka voivat vaikuttaa pelin lopputulokseen.

- Mallien monimutkaisuus ja ylikuormitus

Monimutkaisten ennustemallien käyttö voi olla houkuttelevaa pyrittäessä parantamaan ennustusten tarkkuutta. Kuitenkin liiallinen monimutkaisuus voi johtaa mallin ylikuormitukseen ja liialliseen sovittamiseen historialliseen dataan, mikä voi heikentää mallin yleistettävyyttä ja tehokkuutta uusien tapahtumien ennustamisessa.

- Etiikka ja vastuullisuus

Urheiluedonlyönnin data-analytiikan käytössä nousee esiin myös eettisiä ja vastuullisuuskysymyksiä. Esimerkiksi pelaajien ja joukkueiden yksityisyyden suojaaminen ja reilun pelin periaatteiden noudattaminen ovat tärkeitä näkökohtia, jotka on otettava huomioon data-analytiikan soveltamisessa urheiluedonlyönnissä.

8 PÄÄTÖKSENTEON PARANTAMINEN

Data-analytiikan avulla tehdyt päätökset voivat olla perusteltuja ja ennustettavissa olevia. Urheiluedonlyöjät voivat käyttää näitä analyysseja strategioidensa optimoimiseen ja riskien hallintaan. Se ei kuitenkaan korvaa kokemusta ja tietämystä, vaan täydentää niitä. (Soni, 2023)

Urheiluotteluiden monimutkaiset muuttujat ja vuorovaikutukset tarjoavat haasteita data-analytiikan hyödyntämiselle urheiluedonlyönnissä. Joukkueiden kokoonpanot, pelipaikka, sääolosuhteet ja valmentajien taktiikat ovat vain muutamia esimerkkejä tekijöistä, jotka voivat vaikuttaa pelin dynamiikkaan ja lopputulokseen. Lisäksi urheiluedonlyöntiin liittyy aina ennustamattomuutta ja epävarmuutta, joka voi vaikeuttaa päätöksentekoa. Vaikka data-analytiikka voi auttaa tunnistamaan trendejä ja malleja, se ei voi täysin poistaa yllätyksiä. Loukkaantumiset, epätavalliset tapahtumat ja odottamattomat suoritukset voivat kaikki vaikuttaa pelin lopputulokseen ja haastaa data-analytiikan tehokkuuden vedonlyöntistrategiassa. (Lammey, 2024)

Näiden haasteiden tunnistaminen ja niihin vastaaminen on tärkeää, kun pyritään hyödyntämään data-analytiikkaa urheiluedonlyönnissä. Menestyksellä lähestymistapa vaatii monipuolisen datan keräämistä, jatkuvaa mallien päivittämistä ja tarkkaa riskienhallintaa. Lisäksi on tärkeää ymmärtää, että data-analytiikka on vain yksi osa laajempaa vedonlyöntistrategiaa, joka vaatii myös intuitiota, kokemusta ja markkinoiden tuntemusta.

8.1 Riskien hallinta

Analysoimalla historiallisia tietoja ja tilastoja vedonlyöjät voivat arvioida riskiä eri vedoille ja optimoida vedonlyöntistrategiansa vähentämään tappioita ja maksimoimaan voittoja pitkällä aikavälillä. (Walter, 2023)

Urheiluviedonlyönnissä panosten suuruuden määrittäminen on kriittinen osa riskienhallintaa ja voittomahdollisuuksien maksimoimista. Data-analytiikka tarjoaa tehokkaita työkaluja tämän tehtävän suorittamiseen. Yksi hyödyllisimmistä malleista on Kellyn kaava, joka auttaa vedonlyöjiä löytämään optimaalisen panoskoon jokaiselle vedolle. (Walsh & Joshi, 2024)

Kellyn kaava on kaava, joka laskee optimaalisen panoskoon vedolle odotetun voiton maksimoimiseksi samalla kun minimoi riskin. Se ottaa huomioon vedon todennäköisyyden, kertoimen ja vedonlyöjän pelikassan. Kellyn kaavaa käytetään yleisesti vedonlyönnissä, osakemarkkinoilla ja muilla sijoitusalueilla, joissa on tarjolla tilaisuuksia voittaa tai menettää rahaa. (Kenton, 2022)

$$P = \frac{f - b}{b + 1} \times B$$

Missä:

- P = Panos
- f = Vedon odotettu arvo (kerroin * todennäköisyys - 1)
- b = Vedon todennäköisyys
- B = Vedonlyöjän pelikassa

Alla on esimerkki Python-koodista, joka toteuttaa Kellyn kaavan ja käyttää sitä optimaalisen panoksen laskemiseen urheiluviedonlyönnissä:

```
def kelly_stake(bankroll, odds, edge):  
    """Laskee optimaalisen panoksen Kellyn kaavalla.
```

```

Args:
    bankroll: Vedonlyöjän pelikassa.
    odds: Vedon kerroin.
    edge: Vedon todennäköisyysetu (vedon todennäköisyys -
1).

Returns:
    Optimaalinen panos.
"""
if edge <= 0:
    raise ValueError("Todennäköisyysetu on oltava posi-
tiivinen.")
return (edge * odds) / (odds + 1) * bankroll

# Esimerkki

bankroll = 1000 # Vedonlyöjän pelikassa
odds = 2.0 # Vedon kerroin
edge = 0.05 # Vedon todennäköisyysetu (0.05 = 5%)

optimal_stake = kelly_stake(bankroll, odds, edge)
print("Optimaalinen panos:", optimal_stake)

```

Ohjelma 2. Optimaalisen panoksen laskeminen Kellyn kaavalla. (Kyleskom, 2023)

Tämä koodiesimerkki havainnollistaa, miten Kellyn kaavaa voidaan käyttää Pythonilla optimaalisen panoksen laskemiseen urheiluviedonlyönnissä. Funktio "kelly_stake" laskee panoksen käyttäen annettuja parametreja ja palauttaa optimaalisen panoksen. Lisäksi esimerkki antaa konkreettisen sovelluksen Kellyn kaavan käytöstä urheiluviedonlyönnissä.

On kuitenkin tärkeää huomata, että Kellyn kaavan heikkouksia on se, että siinä todennäköisyydet oletetaan tiedettävän tarkasti ja usein voittotodennäköisyydet ovat viedonlyönnissä vain arvioita. Kellyn kaava ei ota huomioon kaikkia markkinoiden ja sijoittajan yksilöllisiä olosuhteita, kuten riskinsietokykyä tai psykologisia tekijöitä. Monet sijoittajat käyttävät sitä suuntaviivana sijoituspäätöksissään, mutta soveltavat sitä usein harkiten ja muuttavat sen

mukautettuun muotoon ottaakseen huomioon omat tarpeensa ja riskinsietokykynsä.

9 DATA-ANALYTIIKAN YHDISTÄMINEN VEDONLYÖNTI STRATEGIAAN

Tehokas data-analytiikka voi olla merkittävä tekijä NFL-vedonlyönnissä menestyksen saavuttamisessa monella tavalla. Alla on omasta mielestäni parhaimmat tavat, joilla data-analytiikka voi olla avain NFL-vedonlyönnissä:

- Joukkueiden ja pelaajien suorituskyvyn arviointi: Data-analytiikan avulla voi tarkastella historiallisia tilastoja, kuten joukkueiden voittoprosenttia, pisteitä per peli ja yardeja per peli, sekä pelaajien tilastoja, kuten heitto- ja kiinniottojen prosentteja. Näiden tietojen perusteella voi arvioida joukkueiden ja pelaajien vahvuuksia ja heikkouksia eri tilanteissa.
- Trendien seuranta: Seuraamalla trendejä, kuten joukkueiden ja pelaajien viimeaikaista suorituskykyä ja loukkaantumistilannetta, voi tunnistaa kehityskulkuja, jotka voivat vaikuttaa tuleviin otteluihin. Esimerkiksi joukkueen voittoputki tai loukkaantumissuma voi vaikuttaa sen suorituskykyyn seuraavassa pelissä.
- Kertoimien analysointi: Vedonlyöntikertoimet heijastavat vedonlyöjien mielipidettä ottelun lopputuloksesta. Data-analytiikka voi auttaa sinua analysoimaan kertoimia ja arvioimaan, ovatko ne yliarvostettuja tai ali-arvostettuja. Esimerkiksi, jos tietty joukkue on saanut korkean kertoimen, mutta analyysi osoittaa, että sen voittomahdollisuudet ovat todellisuudessa suuremmat, se voi olla hyvä vedonlyöntikohde.
- Simulaatiot ja mallinnus: Käyttämällä matemaattisia malleja ja simulaatioita voit arvioida erilaisten pelitapahtumien todennäköisyyksiä ja luoda strategioita sen perusteella. Esimerkiksi Monte Carlo -simulaatiot voivat

auttaa arvioimaan erilaisten pelitilanteiden todennäköisyyksiä ja tekemään sen perusteella päätöksiä vedonlyönnistä.

Kun rakennat vedonlyöntistrategiaa, on tärkeää yhdistää data-analytiikka ja oma ymmärryksesi NFL:stä. Tässä muutamia vaiheita, joita voi ottaa huomioon strategiasi rakentamisessa:

- Määrittele tavoitteesi: Päätä, kuinka paljon haluat voittaa ja millaisia riskejä olet valmis ottamaan. Tavoitteesi auttavat sinua määrittämään oikean vedonlyöntistrategian.
- Tutki dataa: Kerää ja analysoi laajasti dataa eri lähteistä, kuten tilastollisista tietokannoista, pelivideoista ja asiantuntijoiden ennusteista.
- Kehitä strategia: Valitse strategia, joka sopii tavoitteisiisi ja hyödyntää data-analytiikkaa parhaalla mahdollisella tavalla. Voit esimerkiksi keskittyä arvovedonlyönteihin, hedge-vetoihin tai erilaisiin pelimarkkinoihin.
- Testaa ja hienosäädä: Testaa strategiaasi pienillä panoksilla ja seuraa sen suorituskykyä. Tarvittaessa tee muutoksia ja hienosäädä strategiaasi vastaamaan muuttuvia olosuhteita ja uutta dataa.
- Hallitse riskit: Käytä hyväksi riskinhallintatekniikoita, kuten panoksen kokoa rajoittavia sääntöjä ja hajautettua sijoitusstrategiaa, minimoidaksesi tappioita ja optimoidaksesi voittoja.

Yhdistämällä tehokkaasti data-analytiikka ja vedonlyöntiosaamista voi luoda kilpailukykyisen ja kannattavan vedonlyöntistrategian NFL:ssä.

10 LOPUKSI

Data-analytiikka on tehnyt urheiluedonlyönnistä tarkempaa ja kannattavampaa. Sitä voidaan käyttää monenlaisiin tarkoituksiin, kuten joukkueiden ja pelaajien suorituskyvyn analysointiin, ottelun olosuhteiden arviointiin ja arvokohteiden löytämiseen.

Data-analytiikkaa käyttävät urheiluedonlyönnin sovellukset tarjoavat käyttäjille pääsyn laajaan tietokantaan urheiludataa, jota voidaan käyttää ennusteiden tekemiseen. Sovellukset käyttävät erilaisia data-analytiikkatekniikoita, kuten tilastollista analyysia ja koneoppimista, tehdäkseen ennusteita.

Vedonlyönnissä on aina riski, ja on tärkeää ottaa riskejä harkiten. Urheiluedonlyönnin tulisi asettaa budjetti ja noudattaa sitä. On myös todella tärkeää huomata, että urheiluedonlyönnissä data-analytiikkaa apunaan käyttäminen ei takaa voittoja. Data-analytiikan avulla he voivat kuitenkin parantaa mahdollisuuksiaan menestyä vedonlyönnissä.

Koneoppimismallien käyttö urheilutapahtumien ja -tulosten ennustamiseen tulee todennäköisesti jatkumaan sekä akateemisilla aloilla että kaupallisiin tarkoituksiin. Mahdollisia tulevia tutkimuksia voitaisiin tehdä, miten data-analytiikkaa voidaan soveltaa live-vedonlyönnissä NFL:ssä. Analysoimalla reaaliaikaista dataa ja kehittämällä malleja, jotka voivat auttaa ennustamaan ottelun kulkua ja mahdollisia muutoksia.

LÄHTEET

Altmann, A. (2004). A Statistical Approach to Sports Betting. [väitöskirja, City University London] <https://openaccess.city.ac.uk/id/eprint/8431/>

Bakirov, B., Hodges, Majumdar, A., Scott, S. & Rees, T. (2022) Machine Learning for Understanding and Predicting Injuries in Football.

<https://doi.org/10.1186/s40798-022-00465-4>

Bernard, M. (2015) Big data: using smart bigdata, analytics and metrics to make better decisions and improve performance. <https://ebookcentral.proquest.com/lib/samk/detail.action?docID=1895824>

Bookkerit (n.d) Vedonlyöntiopas: Vedonlyönti vinkit. Haettu 2.3.2024

<https://bookkerit.com/vedonlyonti-vinkit/>

Borghesi, R. (13.6.2007) The home team weather advantage and biases in the NFL betting market <https://doi.org/10.1016/j.jeconbus.2006.09.001>

Borghesi, R. (11.4.2011) The late-season bias: explaining the NFL's home-underdog effect. <https://doi.org/10.1080/00036840600690314>

Brown, L., Jarmoszko, T., Labeledz Jr., C., Schumaker, R. (2017) Prediction from regional angst – A study of NFL sentiment in Twitter using technical stock market charting <https://doi.org/10.1016/j.dss.2017.04.010>

Databasecamp (8.1.2022) What is a Random Forest.

<https://databasecamp.de/en/ml/random-forests>

Degerlund, T. Amerikkalainen jalkapallo: | Näitä ominaisuuksia jenkkifutarilta vaaditaan – tämän jutun luettuasi tunnet amerikkalaisen jalkapallon taktiikan ja pelinrakentajan roolin

<https://www.uusimaa.fi/urheilu/3461692>

ESPN (2024) Super Bowl LVIII Team Stats.

https://www.espn.com/nfl/matchup/_/gameId/401547378

Harrison, C.K. & Bukstein, S. (2016) Analytics informs. Analytics in Action: The sport business analytics process. [Verkkodokumentti].

<https://doi.org/10.1287/LYTX.2016.06.06>

Itotia, S. (2023). The Role of Big Data in Sports Betting: How Analytics Are Shaping Wagering Strategies.techweez.

<https://techweez.com/2023/07/10/role-big-data-sports-betting-analytics/>

Iordanova, T. (22.5.2022) Bet Smarter With the Monte Carlo Simulation.

https://www.investopedia.com/articles/07/monte_carlo_intro.asp

Jaiswal, S. (6.8.2023) The Role of Data Analytics in Sports Betting. Medium.

<https://medium.com/@mycricketbet758/the-role-of-data-analytics-in-sports-betting-61c0fef3ed82>

Jdashbrock (13.3.2023) Sklearn logistic regression example in sports.

<https://thedatajocks.com/sklearn-logistic-regression-example/>

Kenton, W. (02.11.2023) Monte Carlo Simulation: What It Is, History, How It Works, and 4 Key Steps.

<https://www.investopedia.com/terms/m/montecarlosimulation.asp>

Kenton, W. (9.5.2022) Kelly Criterion: Definition, How Formula Works, History, and Goals.

<https://www.investopedia.com/terms/k/kellycriterion.asp>

Krueger, E. (30.8.2021) Data Driven Sports Betting

<https://medium.com/analytics-vidhya/data-driven-sports-betting-3379131a57c>

Kylsekom (17.11.2023) NBA-Machine-Learning-Sports-Betting. Github.

<https://github.com/kyleskom/NBA-Machine-Learning-Sports-Betting>

Lammey, M. (29.1.2024) How to Bet on NFL – The complete guide for 2024.

<https://www.si.com/fannation/betting/how-to-bet/nfl>

Nelson, D. (23.8.2020) Mitä tukivektorikoneet ovat?

<https://www.unite.ai/fi/mit%C3%A4-ovat-tukivektorikoneet/>

News Flash (3.3.2024) The Impact of Data Analytics on Sports Betting Strategies.

<https://medium.com/@newsgroup/the-impact-of-data-analytics-on-sports-betting-strategies-d1be33873cc8>

Nyblom, J. (2015) Yleistetyt lineaariset mallit.

<http://users.jyu.fi/~junyblom/JTMprujub.pdf>

Odds Matrix, (7.7.2023) Sports Betting Analytics: A Key Tool For Bookies To Improve Odds And Boost Profits. <https://oddsmatrix.com/sports-betting-analytics/>

Pempus, B. & Platana, D. (6.2.2024) Forbes: How Does Sports Betting Work?

<https://www.forbes.com/betting/guide/how-does-sports-betting-work/>

Peoples, A. (23.4.2024) Playbook Projections: Implementing Logistic Regression to Predict NFL Game Outcomes.

<https://medium.com/@ahlon.aj.peoples/playbook-projections-model-evaluation-and-verification-c5ef7bb5af2f>

Runkler, T. (2016) Data Analytics. Models and Algorithms for Intelligent Data Analysis. [E-kirja] Haettu: 25.3.2024.1

https://www.academia.edu/43684896/Data_Analytics_Models_and_Algorithms

Silverman, N. & Suchard, M. (2013) Predicting horse race winners through a regularized conditional logistic regression with frailty.

<https://doi.org/10.5750/jpm.v7i1.595>

Solar, D. (28.8.2015) Sports Betting 101: Everything You Need to Get Started.

<https://www.sportsinsights.com/blog/sports-betting-101-everything-need-get-started/>

Soni, V. (27.4.2023) Impact of Big Data on Sports Betting: How Analytics is Changing the Game.

https://www.linkedin.com/pulse/impact-big-data-sports-betting-how-analytics-changing-vishakha-soni/?trk=article-ssr-frontend-pulse_more-articles_related-content-card

The Investopedia Team (22.4.2024) Data Analytics: What It Is, How It's Used, and 4 Basic Techniques.

<https://www.investopedia.com/terms/d/data-analytics.asp>

Walsh, C. & Alok, J. (1.2.2024). Machine learning for sports betting: should model selection be based on accuracy or calibration?

<https://arxiv.org/abs/2303.06021>

Walter (18.8.2023) The Role Of Statistical Analysis In Sports Betting.

<https://thelovelyconcept.com/the-role-of-statistical-analysis-in-sports-betting/>