Tampereen yliopisto

Topias Tiilikainen

TEKOÄLY OSANA HUIPPU-URHEILUN DATA-ANALYTIIKKAA

Informaatioteknologian ja viestinnän tiedekunta
Kandidaattitutkielma
Toukokuu 2023

TIIVISTELMÄ

Topias Tiilikainen: Tekoäly osana huippu-urheilun data-analytiikkaa
Kandidaattitutkielma
Tampereen yliopisto
Tietojenkäsittelytieteiden tutkinto-ohjelma
Toukokuu 2023

Tekoälyn rooli kasvaa läpi yhteiskunnan merkittävää vauhtia, eikä sen käytön kasvu näytä hidastuvan. Tekoälyä käyttäviä ohjelmistoja tai laitteita on hyödynnetty osana huippu-urheilutoimintaa jo useita vuosia ja niiden avulla kerätyn datan avulla joukkueet ja yksilöurheilijat pyrkivät saavuttamaan kilpailuetua. Tässä kandidaatintyössä aiheena on tarkastella tekoälyn käyttöä osana huippu-urheilun data-analytiikkaa. Tutkielmassa käsitellään sitä, miten tekoälyä sovelletaan eri urheilulajeissa ja tutkitaan, miten se toimii apuna suorituskyvyn ja joukkueen ottelustrategian parantamisessa.

Tutkielman keskeiset havainnot ovat, että tekoälyllä on nykypäivänä merkittävä rooli huippu-urheilun data-analytiikassa ja sen käyttö voi tuoda sekä yksilölle sekä joukkueelle monia etuja. Tekoälyn käyttö mahdollistaa tarkemman ja erityisesti nopeamman datan analysoinnin ja sen hyödyntäminen tarjoa uusia tapoja tutkia urheilusuorituksia. Se myös auttaa valmentajia ja urheilijoita tekemään parempia päätöksiä perustuen objektiiviseen, dataan pohjautuvaan analyysiin ja sen pohjalta muodostettuihin ennusteisiin suorituskyvyn osalta.

Tutkielman lopuksi esitetään joitakin haasteita ja rajoituksia tekoälyn käytölle urheiludatan analysoinnissa. Näitä ovat esimerkiksi datan laatu ja määrä sekä taloudelliset resurssit. Tekoälyn hyödyntäminen osana urheilun data-analytiikkaa vaatii aina laitteistoa ja osaamista. Etenkin pienemmissä urheilulajeissa myös huipputasolla on tavallista, että taloudelliset resurssit ovat rajatut. Datan hankinta vaatii myös taloudellisia resursseja tai laitteistoja ja teknologioita, eikä välttämättä siltikään data ole halutunlaista tai tarpeeksi laadukasta analyysiä varten. Lisäksi urheilutoiminnan parista ei välttämättä löydy sisäistä osaamista analyysin harjoittamista varten, jolloin se tulisi ostaa palveluna.

Avainsanat: tekoäly, data-analytiikka, huippu-urheilu, urheiludata, koneoppiminen, urheiluanalytiikka

Tämän julkaisun alkuperäisyys on tarkastettu Turnitin OriginalityCheck –ohjelmalla.

SISÄLLYSLUETTELO

1 Johdanto 1

2 Tutkimusmenetelmät 2

2.1 Tavoitteet ja tutkimuskysymykset 2

2.2 Tutkimuksen toteutus ja ennakoidut tulokset 2

3 Käsitteet ja tieteellinen tausta 3

3.1 Tekoäly 3

3.2 Data-analytiikka 4

3.3 Urheiluanalytiikka 4

3.4 Datankeräysmenetelmät 5

3.4.1 Datankeruulaitteistot 6

3.4.2 Automaattiset datankeruujärjestelmät 6

4 Tekoälyn hyödyntäminen osana huippu-urheilutoimintaa 7

4.1 Jääkiekko 8

4.2 Jalkapallo 10

4.3 Koripallo 11

4.4 Yksilölajit 12

5 Pohdinta 12

6 Yhteenveto 13

Lähdeluettelo 14

1 Johdanto

Tekoäly on yksi maailman keskeisimmistä tulevaisuuden teknologioista ja se vaikuttaa jo tällä hetkellä ihmisten arkeen laajalti, vaikkeivat he sitä itse edes huomaisi. Tekoälyä hyödyntäviä sovelluksia on ollut olemassa jo vuosikymmeniä, mutta etenkin teknologian kehitys, tiedon määrän valtava kasvu sekä uudet algoritmit ovat johtaneet useampiin läpimurtoihin tekoälyn osalta. (Euroopan parlamentti, 2020)

Tekoälyjärjestelmät oppivat ja ovat adaptiivisia, ja yleensä niitä hyödynnetäänkin, kun asiaa ei saada ratkaistua tietokoneita tai ohjelmistoja käyttäen. Tekoälyn soveltamisen lopputuloksena on yleensä kokoelma useampia jo aiemmin kehitettyjä menetelmiä tai algoritmeja. Siltikään tämänhetkiset, tekoälyä käyttävät sovellukset ja laitteistot, eivät älykkyydeltään ole täysin verrattavissa ihmisälyyn, sillä ne eivät omaa ihmismäistä tietoisuutta, tahtoa eikä ymmärrystä oman toimialansa lisäksi. (Heikkinen & Nieminen, 2017)

Tekoälyn hyödyntäminen osana urheilutoimintaa on viimeisten vuosien aikana yleistynyt rajusti (Martin, 2016). Tekoälyä voidaan käyttää huippu-urheilussa esimerkiksi pelaajien tai joukkueen suoritusten tarkkailuun ja analysointiin (Brefeld ym., 2019). Markkinoille kumpuaa jatkuvasti uusia toimijoita uusien tekoälyä hyödyntävien sovellusten tai laitteistojen kanssa (Schumaker ym., 2010). Data-analyysin hyödyntäminen osana taktista lajianalyysiä yleistyy jatkuvasti ja sitä voidaan pitää jopa nykypäivän trendinä (Miller, 2015).

Tekoälyn soveltaminen osaksi urheilutoiminnan analysointiprosessia tarjoaa mahdollisesti uusia näkökulmia ja ajatuksia. Tekoäly kykenee parhaimmillaan esimerkiksi havaitsemaan harjoitus- tai ottelutapahtumista kerätystä datasta malleja, jotka kasvattavat loukkaantumisriskiä. Näin ollen voidaan löytää keinoja ehkäistä loukkaantumisia. (Basole & Saupe, 2016, s. 24; Bernard, 2016)

Tekoälyä hyödyntävien laitteistojen ja ohjelmistojen käyttö osana urheilutoiminnan analyysiprosessia mahdollistaa automaattisten datankeräysmenetelmien hyödyntämisen, laajemman otteluiden analysoinnin sekä tulevaisuuden suorituskyvyn ennustamiseen Lisäksi näiden laitteiden ja ohjelmistojen hyödyntäminen auttaa mahdollisesti tunnistamaan vielä havaitsemattomia trendejä historialliseen dataan perustuen. (Brefeld ym., 2019)

2 Tutkimusmenetelmät

2.1 Tavoitteet ja tutkimuskysymykset

Tutkielman päätavoitteena on selvittää minkälaisilla tekoälyä hyödyntävillä menetelmillä huippu-urheilujoukkueet ovat lähihistorian aikana saavuttaneet kilpailuetua muihin joukkueisiin verrattuna. Tätä tavoitetta selvitetään vastaamalla seuraaviin kysymyksiin: Millaisia tekoälyä hyödyntäviä datankeräysmenetelmiä urheilujoukkueilla on käytössä? Millaisia tekoälyä hyödyntäviä teknologioita urheilujoukkueilla on käytössään? Miten tekoälyjohteisesti kerättyä dataa hyödynnetään osana yksilön tai joukkueen lajianalyysiprosessia? Työn tarkoituksena onkin tutkia, millaisin eri tavoin ja menetelmin tekoälyä hyödynnetään osana urheilujoukkueen tai yksilön toimintaan kohdistuvaa, dataan pohjautuvaa lajianalyysiä, erityisesti huippu-urheilun osalta.

2.2 Tutkimuksen toteutus ja ennakoidut tulokset

Tutkielma toteutetaan kirjallisuuskatsauksena aiempiin tutkimuksiin pohjautuen. Tutkimuksessa käytetty aineisto kerätään käyttäen sähköisiä metodeja. Tutkimusaineistoina hyödynnetään tieteellisiä tutkimuksia, artikkeleita, sekä kirjoja sisältäviä tietokantoja, kuten Andor ja Google Scholar. Tutkimusaineistoa kerätessä pidetään keskiössä tutkimuskysymykset ja analysoidaan löydetyn aineiston soveltuvuutta suhteessa siihen. Aineistoa valittaessa kiinnitetään myös huomiota aineiston pätevyyteen, sillä tarkasteltava aihepiiri on yleisesti nopeasti kehittyvä, joten myös aineiston tulisi olla kohtuullisen tuoretta.

Tutkimuksen näkökulma on tekninen ja tutkimuksessa keskitytäänkin tarkastelemaan aihealuetta aiempien tutkimusten pohjalta. Tutkimuksessa aihealuetta käsitellään yksittäisen huippu-urheilijan tai -urheilujoukkueen näkökulmasta ja tutkitaan tekoälyjohteisen datan roolia osana sen suorituskyvyn kehitystä.

Jo ennen tutkimuksen aloittamista oli tiedossa, että tekoälyn avulla toteutetun dataanalyysin rooli etenkin huippu-urheilussa on kasvanut viimeisten vuosien aikana (Martin, 2016), mutta varsinaiset käytännön tavat soveltaa tekoälyä ovat pitkälti kysymysmerkkejä. Tutkimusta tehdessä ja tarkasteltaessa on hyvä tiedostaa, että tutkimusaineistossa haasteeksi saattaa muodostua urheilujoukkueiden saavuttaman kilpailuedun varjelu, jolloin seikkaperäisiä esimerkkejä käytetyistä sovelluksista ja saavutetuista hyödyistä voi olla hankala löytää. Avoimesti saatavilla olevaa dataa on kuitenkin valtaosaan lajeista

saatavilla edes jonkun verran, ja näitä onkin havainnollistettu jossain määrin myös tutkielman tekoa varten kerätyssä aineistossa.

3 Käsitteet ja tieteellinen tausta

Tässä luvussa esitellään työssä esiintyvä termistö ja niiden tarkoitusperä niiltä osin kuin se tutkimuksen lukemisen kannalta on aiheellista.

3.1 Tekoäly

Tekoäly on käsitteenä hankalasti määriteltävissä, eikä sitä voida rajata yhden tietyn määritelmän sisälle. Se sisältää monia eri menetelmiä, teknologioita, sovelluksia ja liittyy useisiin eri tieteenaloihin. Lisäksi se vaikuttaa jossain määrin koko yhteiskuntaan ja ihmisiin teknologisen kehityksen lisäksi. Tekoälyn muodot on yleensä määritelty jaettavaksi kahteen eri kategoriaan: kapeaan, eli heikkoon tekoälyyn, sekä vahvaan tekoälyyn. Kaikki nykypäivän tekoälyä hyödyntävät sovellukset on luokiteltu kapean tekoälyn alaiseksi, sillä vahvaksi tekoälyksi mielletään täysin ihmisen kaltaisen tietoisuuden ja ymmärryksen omaavaa tekoälyä. (Heikkinen & Nieminen, 2017)

Euroopan parlamentti (2020) määrittelee nettisivuillansa tekoälyn seuraavasti: ”Tekoälyllä tarkoitetaan koneen kykyä käyttää perinteisesti ihmisen älyyn liitettyjä taitoja, kuten päättelyä, oppimista, suunnittelemista tai luomista”. He myös lisäävät, että tekoälyä on vuosien saatossa hyödynnetty laajasti osana eri tieteenalojen sovelluksia ja teknologioita. Euroopan parlamentin mukaan tekoälyä hyödynnetään usein ihmisille tarjottujen palveluiden yhteydessä. Heidän määritelmänsä mukaan tekoälyn avulla pystytään kohdentamaan ja optimoimaan palveluita sopivammaksi ihmisille, sekä esimerkiksi parantamaan tehokkuutta ja turvallisuutta. Tekstissä myös mainitaan, että tekoälyä hyödynnetään usein esimerkiksi ennakoimaan mahdollisia luonnonkatastrofeja, ehkäisemään disinformaation levittämistä tai vaikkapa apuna kielen käännöksissä.

Brown (2021) kertoo artikkelissaan, että yleisimpiä käsitteitä tekoälyn yhteydessä ovat koneoppiminen ja neuroverkot. Artikkelin mukaan koneoppimista käyttävä järjestelmä tai ohjelma oppii syötetyn datan ja dataa prosessoivien algoritmien avulla, sekä että neuroverkot ovat yksi koneoppimisen alalaji. Lisäksi hänen mukaansa neuroverkot pyrkivät mukaan jäljittelemään ihmisaivojen toimintaa liikuttaen dataa sen sisältämiä liitoksia pitkin eri funktioita ketjuttaen.

3.2 Data-analytiikka

Data-analyysistä tuli suosittua heti 2000-luvun alkutaipaleella (Kohavi ym., 2002, s. 45-48; Tyagi, 2003, s. 12-14). Data-analyysilla tarkoitetaan yleisesti suurten tietojoukkojen analysointia päätöksenteon tueksi hyödyntäen tietokonejärjestelmiä ja -sovelluksia. Alana data-analyysi on hyvin laaja-alainen ja monitieteinen ja sillä on piirteitä monilta muilta tieteenaloilta. Isona osana data-analyysiä voidaan pitää esimerkiksi tilastollista mallintamista ja tilastotiedettä yleisesti, koneoppimista sekä tekoälyä. Data-analyysin päätarkoituksena on löytää mielenkiintoisia ja hyödyllisiä seikkoja datasta, sekä esimerkiksi toistuvia malleja eli trendejä. (Runkler, 2016).

Runklerin (2016) mukaan data-analyysiprosessi on yleensä jaettavissa myös useampaan eri vaiheeseen tai osa-alueeseen. Hän kertoo, että data-analyysiprosessi sisältää tyypillisesti datan keräämisen ja valmistelemisen, sen esiprosessoinnin, eli kerätyn datan puhdistus ja suodatus, sekä analysoinnin ja visualisoinnin, jonka jälkeen analyysitulokset vielä lopuksi arvioidaan ja tulkitaan.

3.3 Urheiluanalytiikka

Alamarin (2013, s. 5–6) määritelmän mukaan urheiluanalytiikka on historiallisen ja rakenteellisen jatkuvasti kasvavan datan hallinnointia, ennakoivien tilastollisten mallien soveltamista kyseiseen dataan, sekä erilaisten järjestelmien hyödyntämistä datan käyttämiseksi osana päätöksentekoprosessia. Davenport (2014) taas määrittelee sen olevan data-analytiikan täsmällistä hyödyntämistä ammatillisen urheilun saralla.

Brefeld ja muut (2019) toteavat tutkimuksessaan, että urheilussa joukkueet pyrkivät datan avulla ehkäisemään loukkaantumisia, analysoimaan ottelutapahtumia, luomaan ja muokkaamaan strategioita ottelu- ja harjoittelutapahtumiin sekä parantamaan ja ennustamaan tulevaisuuden suorituskykyä. Myös Bernard (2016) yhtyy näkemykseen, ja hänen mukaansa dataa voidaan hyödyntää urheilussa optimoimaan yksilön unta, palautumista ja harjoittelua. Macdonald (2012) tuo tutkimuksessaan ilmi, että urheiluanalytiikan tavoitteena on pelaajien ja joukkueiden suorituksen analysointi mahdollisimman objektiivisin silmin. Barlow (2015) toteaakin, että etenkin joukkueurheilussa päätöksiä tekevät henkilöt käyttävät entistä enemmän aikaa datan keräämiseen ja sen analysointiin.

Schumaker ja muut (2010) luokittelevat urheilussa tiedontarpeen viiteen eri luokkaan. Heidän tekemänsä luokittelu perustuu lineaarisesti etenevään portaittaiseen määrittelyyn, joista ensimmäisessä dataa ei juurikaan hyödynnetä ja viidennessä hyödynnetään erilaisia tiedon louhinnan tekniikoita päätöksenteon tukena. Teoksessaan he mainitsevat, että hyvin pieni määrä urheiluorganisaatioita sijoittuu luokittelussa viidenteen luokkaan, ja että datan hyödyntämisellä paremmin kuin muut on mahdollista saavuttaa kilpailuetua.

Kerätyn datan pohjalta on mahdollista luoda erilaisia malleja, jotka tuottavat johtopäätöksiä tai analyyseja ottelutapahtumista tai yksittäisistä tilanteista. Tällaista mallia voidaan esimerkiksi jalkapallossa soveltaa tilanteeseen, jossa halutaan selvittää millä todennäköisyydellä tietystä sektorista tai sijainnista potkaistu potku tuottaa keskimäärin maalin. (Brefeld ym., 2019)

Macdonald (2012) ja Pellinen (2019) esittelevät töissään jääkiekkoon sovellettua laukauksen maaliodottama -mallinnusprosessia, jonka avulla jokaiselle jääkiekko-ottelussa tuotetuille laukauksille pystytään arvioimaan keskimääräinen todennäköisyys, jolla laukauksen lopputulemana on maali. Brefeld ja muut (2019) esittelevät vastaavanlaisen mallinnusmenetelmän jalkapalloon sekä muita ennusteita tuottavia mallinnusmenetelmiä, joiden avulla esimerkiksi jalkapallossa, jääkiekossa ja maantiepyöräilyssä pystytään arvioimaan pelaajien suorituksia historiallisen datan perusteella sekä lisäksi luomaan näiden arvioiden pohjalta ennusteita tulevista suorituksista.

Azar ja muut (2018) toteavat, että systemaattisen analyysiprosessin avulla päätöksentekijöillä on mahdollista vahvistaa ottelutapahtumiin perustuvia päätöksiä. Davenport (2014) mainitsee, että kerätyn datan ja siitä johdettujen tietojen perusteella valmentajat voivat luoda ottelusuunnitelmansa objektiivisesti.

3.4 Datankeräysmenetelmät

Nykypäivänä urheilussa, kuten myös yleisesti maailmassa, dataa kerätään ja on saatavilla enemmän kuin sitä ehditään hyödyntämään. Tämä johtunee siitä, että modernien teknologioiden ja työkalujen ansiosta datan keräämisestä on tullut helppoa ja sitä on edullista säilyttää. (Miller, 2015)

Schumakerin ja muiden (2010) mukaan urheiluorganisaatioiden lisäksi tilastointia ottelutapahtumista tuottavat myös ulkoiset toimijat, jotka luovat tilastopohjaisia analyyseja ja jakavat tuotoksensa eteenpäin. He kertovat, että osa toimijoista on kaupallisia toimijoita, mutta tarjolla on myös monia yhteisölähtöisiä palveluita. He tuovatkin ilmi useita eri esimerkkejä eri verkkosivustoista, jotka avoimesti tarjoavat nähtäville laji- ja sarjakohtaista tilastointia. Heidän mukaansa urheiluorganisaatiot pitävät yleensä itse keräämänsä datan omana tietonansa.

Nykypäivänä videoiden, laitteistojen tai muiden tekniikoiden ja ohjelmistoratkaisujen avulla kerättyä dataa on yleensä myös mahdollista tarkastella reaaliajassa yleisesti käytössä olevia laitteistoja tai sovelluksia käyttäen. Useimmiten datankeruulaitteistot ja -menetelmät ovat yhteensopivia useampien älypuhelimien, tietokoneiden ja tablettien kanssa tai omaavat käyttöliittymällä varustetun sovelluksen datan tarkastelua varten. Useimpien välineistöjen osalta data voidaan tarvittaessa myös ladata esimerkiksi pilvipalveluihin myöhempää käyttöä varten. (Barbosa, 2018)

3.4.1 Datankeruulaitteistot

Puettavat mittauslaitteet ovat kehittyneet viime vuosien aikana erittäin nopeasti, sekä niiden käyttö on yleistynyt merkittävästi urheilun saralla (Martin, 2016). Aroganamin ja muiden (2019) mukaan yleisesti käytetyin puettava datankeruulaite on ranteeseen kiinnitettävä sensorilaite, mutta useissa lajeissa, kuten esimerkiksi amerikkalaisessa jalkapallossa, golfissa, jalkapallossa ja jääkiekossa, dataa kerätään osaksi pelaajien varusteita sijoitettujen sensoreiden avulla. He tuovat myös ilmi, että sensorit ja niiden sijoituspaikka pelaajassa tai pelaajan varusteissa vaihtelee yleensä tapauskohtaisesti lajin ja halutun datan osalta. Miller (2015) ja Ebling (2016, s. 2) tuovat ilmi, että puettavien datankeruulaitteiden pääkäyttötarkoitus on suorituksen aikainen kehon toiminnan mittaaminen. Molemmat heistä mainitsevat tällaisiksi mitattavaksi arvoiksi esimerkiksi sykearvot, hengitystiheyden sekä askeltiheyden. Aroganam ja muut (2019) kertovat seurantalaitteistojen hyödyntävän erilaisia menetelmiä ja tekniikoita, kuten kiihtyvyysantureita, gyroskooppeja, magnetometrejä, GPS-järjestelmää ja painesensoreita, joiden avulla laitteet tuottavat esiprosessoitua dataa käyttäjästään ja tuovat sen automaattisesti käyttäjän nähtäville esimerkiksi pilvipalveluja hyödyntäviin sovelluksiin.

Eblingin (2016, s. 2) mukaan datan kerääminen onnistuu myös ei-puettavien, yleensä pelivälineisiin tai varusteisiin asennettavien sensorien tai laitteiden avulla. Hän kertoo, että kyseiset laitteet hyödyntävät toiminnassaan samoja teknologioita kuin useimmat puettavatkin datankeruulaitteet ja pystyvät samalla tapaa tekemään havaintoja mittaamastaan objektista. Yksi esimerkki Eblingin mainitsemista laitteista on Adidaksen kehittämä FIFA:n vuoden 2022 World Cup -lopputurnauksessa virallisena kisapallona käytetty älyjalkapallo, jonka pallon sisään asennettu sensoriverkko kerää 500 datapistettä jokaista sekuntia kohti (Adidas, 2022).

Miller (2015) kuitenkin huomauttaa tutkimuksessaan, että useamman laitteen tai datalähteen yhdistelemisestä syntyvä datamäärä aiheuttaa ongelmia ammattiurheilujoukkueiden keskuudessa teknologian kehityksen kyydissä pysymisen osalta.

3.4.2 Automaattiset datankeruujärjestelmät

Urheilujoukkueiden tai -yksilöiden käytössä on myös automaattisesti toimivia datankeruujärjestelmiä (Linke ym., 2020; Lukan, 2022; Sellitto, 2015; Wisehockey, 2022). Teknologian kehityttyä ja vuosien aikana massiivisesti yleistyneiden kameralaitteistojen myötä mahdollisuus ottelu- tai harjoitustapahtumien videokuvaamiseen ja täten videodatan keräämiseen on eri urheilulajien ja -sarjojen osalta yleistynyt (Aroganam ym., 2019).

Videodata on eri urheilulajien parissa merkittävin kilpailuetua tuova datan tyyppi. Sen hyödyntäminen helpottaa urheilujoukkueiden ja yksilöiden fyysisten sekä taktisten analyysien tuottamista. Videodataa voidaan hyödyntää myös otteluihin valmistautumisessa, mutta haasteena sen käytössä on merkittävien kenttätapahtumien erottelu tallennetulta videolta automaattisesti yksittäisiksi videoleikkeiksi. (D'Orazio & Leo, 2010, s. 2911–2926).

Videokuvan avulla on mahdollista tuottaa automaattisesti tilastoja yksinkertaisista kenttätapahtumista, kuten syöttöjen tai laukauksien määrästä. Monimutkaisempien ja perusteellisempien tilastojen tuottaminen videokuvan avulla on kuitenkin edelleen haastavaa ja vaatii manuaalista työtä. (Davenport, 2014).

Useiden eri lajien huippusarjat ovat myös vuosien aikana panostaneet omiin automaattisiin datankeräysmenetelmiinsä. Esimerkiksi Pohjois-Amerikkaan sijoittuva NBA-koripalloliiga on ottanut käyttöönsä SportVU -järjestelmän, joka kerää dataa automaattisesti pelivälineestä ja jokaisesta pelaajasta yhteensä 25 datapistettä sekunnissa (Sellitto, 2015).

Myös jalkapallossa on hyödynnetty jo vuosien ajan vastaavan kaltaisia järjestelmiä. Esimerkiksi useissa Euroopan suursarjoissa on hyödynnetty ChyronHegon tuottamaa TRACAB-järjestelmää vuodesta 2013 alkaen. TRACAB-järjestelmässä joukkueiden saatavilla on kolmen tyyppistä dataa: reaaliaikaista dataa, noin 15 sekunnin viiveellä tekoälyn avulla prosessoitua hieman reaaliaikaista virheettömämpää dataa, sekä muutaman tunnin päästä tallennuksesta järjestelmän lopullisesti prosessoimaa dataa. TRACAB-järjestelmä hyödyntää videokuvan perusteella suoritettua optista seurantaa ja kykenee määrittämään jalkapallokentän jokaiselle datapisteelle (pelaajat, tuomari sekä peliväline) sijaintitietoja 25 kertaa sekunnissa. (Linke ym., 2020).

Sijaintidataa hyödyntävät teknologiat tuottavat käyttäjälle tietoa esimerkiksi pelaajien liikehdinnästä. Tämä voi sisältää muun muassa tietoa pelaajan toteuttamista kiihdytyksistä, jarrutuksista tai esimerkiksi pelaajien välisten kontaktien määristä. Yleisesti sijaintitietoihin pohjautuvat teknologiat tuottavat myös dataa pelaajakohtaisesti liikutun matkan sekä korkealla intensiteetillä toteutettujen juoksujen määristä. (Davenport, 2014).

4 Tekoälyn hyödyntäminen osana huippu-urheilutoimintaa

Eri teknologioiden ja välineistön avulla kerättyä dataa hyödynnetään laajalti läpi lajirajojen myös harjoitusohjelmien suunnittelun tukena, sekä loukkaantumisten ehkäisyyn ja minimoimiseen. Basolen ja Saupen (2016, s. 24) mukaan, etenkin huipputasolla pelaajille luodaankin yleensä henkilökohtaisia harjoitus- tai palautumisohjelmia perustuen heistä

kerättyyn aktiivisuus- tai sykedataan. Kerätty ja analysoitu data myös kykenee, riippuen mallinnustavasta, mahdollisesti tuomaan ilmi sellaisia seikkoja tai trendejä, joita ihmismieli ei välttämättä pysty ottamaan huomioon (Runkler, 2016).

4.1 Jääkiekko

Liiga (kotimaisen jääkiekon korkein sarjataso) otti kaudesta 2019–2020 alkaen käyttöönsä älykiekkojärjestelmän, jonka avulla Liigan eri sidosryhmien saatavilla on otteluista kerättyä dataa numeromuodossa (Bitwise Oy, 2019). Kyseinen järjestelmä perustuu jääkiekon ja jääkiekkohallien sisälle asennettaviin sensoreihin, jotka suorittavat reaaliaikaista sisäpaikannusteknologiaa. Automaattijärjestelmässä käytettävän sisäpaikannusteknologian kehittäneen Quuppa Oy:n (2019a) mukaan järjestelmä kerää otteluista muun muassa sykelukemia pelaajakohtaisesti ja nämä ovat suoraan joukkueiden hyödynnettävissä. Järjestelmän keräämä data muotoillaan automaattisesti algoritmien ja tekoälyn avulla helpommin ymmärrettäväksi tekstiksi tai grafiikaksi ja joukkueet pystyvät hyödyntämään sitä jo otteluiden aikana (Quuppa Oy, 2019b). Bitwise Oy:n kehittämää Wisehockey -järjestelmää hyödyntävät myös Saksan jääkiekkoliiga DEL, pääosin Venäjän maaperälle sijoittuva KHL, sekä Norjan jääkiekkoliiga Fjordkraft Ligaen (Wisehockey Oy, 2022).

Miten ja mihin tätä kaikkea dataa kerättyä sitten käytetään ja hyödynnetään? Jääkiekon osalta Quuppa Oy (2019a) tuo julkaisemassaan case-study-artikkelissaan esille, että käytetyn teknologian avulla pelaajat ja valmentajat saavat reaaliaikaista, objektiivista, dataa suorituskyvystään. Joukkueilla on myös mahdollisuus tuottaa omia johtopäätöksiä kerätystä datasta omien mallinnusprosessien avulla. Useasti nämä sidosryhmät näkevät ottelun subjektiivisesti, eivätkä välttämättä kykene muodostamaan täysin puolueetonta kuvaa ottelutapahtumista. Quuppa Oy (2019a) lisääkin artikkelissaan, että datasta johdetut analyysit ovat oiva apuväline muodostamaan puolueetonta näkökulmaa ottelun kulusta ja tekemään sen avulla johtopäätöksiä.

Näiden automaattisesti muodostettujen tilastojen ja havaintojen avulla joukkueet kykenevät esimerkiksi tekemään muutoksia ottelustrategiaansa parempaa suorituskykyä tavoitellen (Ryhänen, 2020, s. 22-26). Quuppa Oy:n (2019b) julkaisemassa case-study-videossa näkyvien eri tilastojen avulla esimerkiksi pitkiksi venyneiden vaihtojen ja pelaajien sykearvojen vaikutus ottelutapahtumiin voidaan havainnollistaa lukujen ja graafien muodossa, sillä ne ovat suoraan nähtävissä sovelluksen käyttöliittymässä. Kyseiseltä videolta myös nähdään, että automaattijärjestelmä kerää yksittäisten ottelutapahtumien lisäksi tietoa muun muassa jäällä olevista ketjukombinaatioista ja ottelutapahtumat ovat niiden mukaan luokiteltavissa. Automaattijärjestelmän datapisteistä johdetun datan avulla joukkueiden on siis mahdollista rakentaa erilaisia tilastollisia mallinnusprosesseja

analysoidakseen ottelutapahtumia tilastollisesta näkökulmasta. Näitä voidaan hyödyntää reaaliaikaisuuden myötä jo ottelun sisällä esimerkiksi joukkuestrategian tai pelaajakohtaisen palautteen osalta. Konkreettisia esimerkkejä voisi jääkiekon osalta olla esimerkiksi vaihtopituuksien optimointi suhteessa pelaajien kuntoon. (Quuppa Oy 2019b)

KUVA 1: Wisehockey älykiekkojärjestelmän käyttöliittymä (Bitwise Oy, 2019)

Vastaavan kaltaista järjestelmää hyödyntää datankeruussa ja tilastojen luomisessa myös NHL. Lukan kertoo, että he hyödyntävät järjestelmän keräämää dataa NHL-seura Seattle Krakenissa laajasti ja analysoivat sen avulla kaikkea koko joukkueen ja yksittäisen pelaajan välillä. Hän myös mainitsee, että suoraan yksittäisiä tilastoja lukemalla voi myös harhaantua analyyseissään, sillä vaikka datasta kerätyt tilastot selittävätkin jotain, tulee aina huomioida myös tilaston konteksti ja sen mahdollisesti sisältämä kohina. Lukan kertookin, että tilastoja yleensä skaalataan joko peliminuutteihin tai ottelumääriin suhteutettuna, mikä vähentää luvuissa esiintyvää varianssia ja helpottaa tilastojen vertailua esimerkiksi pelaajien välillä. Hän myös mainitsee, ettei kaikki NHL:n automaattisesti kerättävä data kuitenkaan ole julkisesti saatavilla, jonka vuoksi esimerkiksi Seattle Kraken kerää itse manuaalisesti osan datastansa. He keräävät esimerkiksi tilastointeja syöttöketjuista ja hyödyntävät tätä tietoa toiminnassaan. Näistä heidän keräämistään tilastoista on esimerkiksi mahdollisuus tuottaa analyysejä, jotka kertovat minkälaiset syöttökombinaatiot johtavat todennäköisimmin maaliin, sekä miten joukkue tehokkaimmin pääsee vastustajan hyökkäysalueelle. (Lukan, 2022)

4.2 Jalkapallo

Jalkapallossa yleisimmät datankeruumenetelmät ovat samankaltaisia kuin jääkiekossa. Huipputasolla, kuten esimerkiksi vuoden 2022 syksyn Qatarin MM-lopputurnauksessa, pelaajat varustetaan yleensä televisiokuvissa satunnaisesti vilahtavilla GPS-sykesensoreilla, videokuvasta kerätään automaattisesti datapisteitä tekoälyn avulla, sekä pelivälineenä toimii mahdollisesti älypallo (Adidas 2022; Tiga, 2022).

Lewis (2021) kertoo artikkelissaan, että Englannin Valioliigassa on viimeisten vuosien aikana tehty paljon toimia datan hyödyntämisessä kilpailuedun luomiseksi. Lewiksen mukaan esimerkiksi Manchester City on palkannut riveihinsä muun muassa rakettitieteilijöitä ja Arsenal hyödyntää toiminnassaan ohjusseurantajärjestelmiä. Hän myös kertoo, että valtaosalla Valioliigan joukkueista on palkkalistoillaan analyytikkotiimi, joka tuottaa dataa päätöksentekijöille taukoamatta. Artikkelissa mainitaan lisäksi, että seurat hyödyntävät analyytikkoja laajalti kykyjenetsinnän apuna, ottelu- ja harjoitteluanalyyseissä ja ennakoimaan todennäköisiä tapahtumia otteluiden aikana tiettyjä vastustajia vastaan.

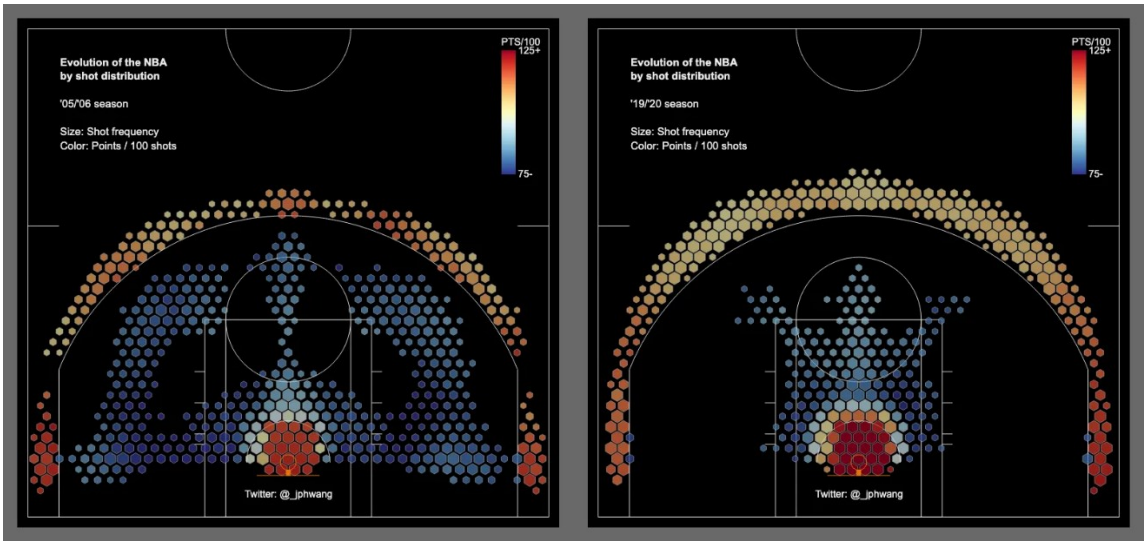
Yksi tunnettu esimerkki data-analyysin hyödyntämisestä jalkapallossa on Englannin Valioliigaa pelaava Brentford FC. Vuonna 2012 urheiluvedonlyöjä Matthew Benham pelasti Brentfordin vararikon partaalta ja aloitti radikaalit dataan perustuvat muutostyöt seuran sisällä. Benham ymmärsi seuran ostettuaan, että taistellakseen pienenä ja rajallisen budjetin omaavana jalkapalloseurana suurempia ja ison budjetin joukkueita vastaan, heidän tulisi toimia muita tehokkaammin. Benham onnistuikin luomaan uuden lähestymistavan, jossa matemaattisia malleja hyödyntämällä onnistuttiin löytämään joukkueen suorituskykyä parantavia pelaajia suhteellisen halvalla siirtohinnalla. Benham myös teki epätavallisia päätöksiä tilastojen pohjalta: hän lopetti seuran oman pelaajatuotannon, sillä kasvatetut nuoret pelaajat houkuteltiin yleensä suurempien seurojen riveihin ilman rahallista korvausta. Seuran suorituskyvyn analysoinnissa olivat keskiössä edistyneet tilastot, kuten maaliodottamaan perustuvat mallit. sekä Näiden pohjalta tehtiin myös havaintoja optimoimaan suorituksia tulevaisuudessa. (Arastey, 2018).

Stanojevic ja Gyarmati (2016) esittelevät tutkimuksessaan otteluista kerättyyn dataan pohjautuvan pelaajan arviointiin soveltuvan mallinnusprosessin, jonka avulla pystytään luomaan arvio pelaajan arvosta hänen historiallisten peliesitystensä avulla. Heidän käyttämänsä malli hyödyntää lukuisia datapisteitä ja kykenee tunnistamaan pelaajia, joihin seurajoukkueiden kykyjenetsijöiden olisi mallin mukaan kannattavaa kohdistaa katseensa tarkemmin. He mainitsevat, että tällainen mallinnusprosessi voi olla urheiluseuroille erittäin arvokas, sillä seurojen seuraamat potentiaaliset hankinnat voivat pelata ympäri maailmaa. Heidän mukaansa tällaisen mallinnusprosessin avulla voidaan analysoida potentiaalisia pelaajia etäältä käsin datan avulla ja siten laittaa paremmuusjärjestykseen.

4.3 Koripallo

Kuten luvussa 3.4.2 todettua, myös Bechtold (2021) vahvistaa, että paikannukseen perustuvia teknologioita sovelletaan koripallossa ainakin lajin kovatasoisimmaksi sarjaksi mielletyn NBA:n osalta.

Abbas (2019) tuo kirjoituksessaan ilmi konkreettisia esimerkkejä, jotka ovat mullistaneet joukkueiden pelitapaa NBA:ssa. Artikkelin mukaan alle vuosikymmen takaperin vain muutamilla NBA-joukkueilla oli palkkalistoillaan data-analyytikko. Nykyään asetelmat ovat päinvastoin. Abbas kirjoittaa, että koripallossa yleisimmät käyttökohteet data-analyysin osalta mukailevat pitkälti jääkiekon ja jalkapallon osalta tehtyjä havaintoja keskittyen kykyjenetsintään ja maaliodottamamallinnukseen tähtäimenään suorituskyvyn parantaminen. Abbas esittää myös data-analyysistä johtuvia konkreettisia esimerkkejä lajin evoluutiosta NBA:n osalta, kuten heiton lähdön sijainnin muuttuminen vuosien saatossa. Hänen mukaansa 2000-luvun alussa yksittäisessä NBA-ottelussa suoritettiin keskimäärin noin 12,5 kolmen pisteen heittoyritystä. Nykyään vastaava luku lähentelee 30 heittoyritystä ottelua kohti. Abbas kertoo syyn olevan tilastoihin pohjautuva data-analyysi, jonka avulla on todettu kolmen pisteen heittoyrityksen olevan kannattava, mikäli heittäjä pussittaa keskimäärin vähintään 33 % kaikista heittämistään kolmen pisteen heittoyrityksistä. Lisäksi Abbaksen mukaan ainakin osa seuroista myös hyödyntää pelaajista automaattisesti kerättyä dataa myös optimoidakseen yksittäisten pelaajien suorituskykyä ottelun sisällä.



KUVA 2: NBA heittoyritysten lämpökarttojen kehitys visualisoituna (Hwang, J. 2020)

Tämä kolmen pisteen heittoyritysten kehitys on havainnollistettu myös Baileyn ja muiden (2018) tutkimuksessa, jossa he tutkivat NBA:ssa tapahtuvien 3-pisteen heittoyri-

tystä edeltävien tapahtumien vaikutusta heiton onnistumisen todennäköisyyteen. He tuovat tutkimuksessaan ilmi mielenkiintoisia tuloksia muun muassa sijainnin ja heittokulman suhteesta heittotodennäköisyyteen, minkä osalta Cheeman (2019) mukaan NBA-joukkueet ovatkin heränneet ja muokanneet taktiikkaansa vuosien saatossa, kuten Abbaskin totesi.

4.4 Yksilölajit

Gain (2022) kertoo, että esimerkiksi tenniksen osalta automaattisia datankeruujärjestelmiä on käytössä ainakin Wimbledonin legendaarisessa turnauksessa. Artikkelissa kerrotaan Wimbledonin järjestävän tahon IBM:n keräävän automaattisesti dataa tennisottelusta niin tilastoihin, kuin tuomaroinnin avuksi. Hänen mukaansa esimerkiksi IBM:n tarjoamat tilastot, eli pelaajien syöttö-, virhe-, ja IBM:n itse kehittämä voimaindeksi, tulevat automaattisesti kerätyksi heidän järjestelmänsä avulla. Kovalchik (2021) kuitenkin tuo ilmi, että avointa dataa tenniksessä on erittäin rajallisesti saatavilla. Hänen mukaansa dataa keräävät yhtiöt, pääasiassa turnausjärjestäjät, keskittyvät lähinnä maksimoimaan liikevoittonsa, eikä esimerkiksi pelaajayhdistystä ole vielä toistaiseksi huolehtimassa pelaajien oikeuksista.

Uinnissa ja hiihdossa datan avulla pyritään optimoimaan lajissa tapahtuvia liikeratoja ja tekniikoita, sekä ehkäisemään loukkaantumisia. Kuten nykypäivänä muuallakin, lajeissa on laajasti käytössä puettavia datankeruulaitteita, joiden keräämää dataa seurataan ja hyödynnetään osana urheilijan toimintaa ja harjoitusohjelman suunnittelua. (Tietoevry., 2021; Vesterinen ym., 2009; Xu, S., & Lee, S., 2022).

5 Pohdinta

Kuten jo luvuissa 3 ja 4 todettiin, datan ja erilaisten mallinnusmenetelmien avulla voidaan eri lajeissa tehdä analyysejä ja johtopäätöksiä näiden analyysien pohjalta, joihin joukkueet voivat pohjata esimerkiksi strategiset muutokset ottelusuunnitelmaansa (Azar ym., 2018). Tehdyistä muutoksista on myös saatavilla nykyaikaisten datankeruumenetelmien ansiosta yleensä liki välitöntä palautetta, sekä taktisten muutosten vaikutuksia pystytään tutkimaan läpi valitun ajanjakson historiaan perustuvan datan pohjalta (Barbosa, 2018). Muutoksien vaikutuksia pystytään myös ainakin jossain määrin ennustamaan luotettavasti datan, mallinnusmenetelmien ja analyysien laadun mukaan (Brefeld ym., 2019).

Haasteita ja rajoitteita datan hyödyntämisen osalta kuitenkin löytyy edelleen. Monissa kilpasarjoissa dataa kerätään ja jaetaan osallistujien käyttöön järjestäjätahon toimesta, mutta datassa voi esiintyä puutteita tai se ei välttämättä ole suoraan urheilijoiden

tarpeisiin soveltuvaa. Dataa ei myöskään välttämättä yksinkertaista tilastointia enempää ole saatavilla, jolloin soveltaakseen data-analyysiä, se on kerättävä itse. Datan laatu ja määrä voi myös olla jossain tapauksissa puutteellinen, jolloin sen kerääminen tai käsitteleminen vaatii manuaalista työtä. Kuten aiemmin mainittua, dataa ja eri datankeruujärjestelmiä ja -laitteistoja voi olla myös liikaa käytössä, jolloin ne eivät välttämättä ole yhteensopivia keskenään. Liika datan määrä itsessään voi myös tuoda omat haasteensa (Miller, 2015), joten dataa hyödyntävien tulisi osata tunnistaa sen joukosta oman toimintansa osalta olennaiset seikat. Tulee myös muistaa, että laajamittaista data-analyysiä harjoitettaessa vaadittujen resurssien määrä kasvaa merkittävästi, eikä yksittäisillä urheilijoilla tai pienillä urheiluorganisaatioilla välttämättä ole riittävää kapasiteettia vastata vaateisiin. Usein etenkin pienempien lajien osalta taloudelliset resurssit ovat varsin rajalliset, ja datan hyödyntäminen vaatii kuitenkin yleensä ymmärrystä ja osaamista, tai palvelun hankintaa muualta.

Myös datan keruuseen ja hallintaan liittyvät yksityisyysseikat ovat mielenkiintoisia. Esimerkiksi tenniksen osalta tutkimuksessa nousi esille kysymys siitä, että ovatko pelaajat oikeutettuja saamaan haltuunsa dataa heidän otteluistaan ollessaan pelissä osallisena. Toistaiseksi näin ei ole ollut, mutta tulevaisuudessa tilanne voi muuttua, mikäli tennikseen saadaan pelaajayhdistys ja neuvottelut asiasta käynnistyvät. Tämä voi avata uusia ovia datan hyödyntämisessä, kun dataa ei enää käytetä pelkästään viihdearvona ottelulähetyksissä tai tuomaroinnin apuna.

6 Yhteenveto

Toteutettu tutkimus osoitti, että etenkin huippu-urheilussa on laajalti käytössä yleensä useampia datankeräysmenetelmiä ja teknologioita (Aroganam ym., 2019). Tekoälyjohteisesti kerättyä dataa hyödynnetään yksilöiden ja joukkueen analyysiprosessissa tavoitellen parempaa suorituskykyä. Käytetyt ohjelmat ja laitteistot myös auttavat tunnistamaan suorituskyvyn kannalta merkittäviä tekijöitä, kuten fyysisiä ominaisuuksia, sekä antamaan suosituksia joukkueen lajitaktiikan osalta (Aroganam ym., 2019; Brefeld ym., 2019; Runkler, 2016).

Seikkaperäisiä esimerkkejä saavutetuista hyödyistä huippu-urheilun osalta on vaikea löytää, sillä kuten jo tutkimuksen alussa aavisteltiin, joukkueet ja yksilöt haluavat todennäköisesti pitää kiinni mahdollisesti saavutetusta kilpailuedustaan. Tutkimuksen yhteydessä kuitenkin löydettiin muutamia ihan seikkaperäisesti esitettyjä mallinnusmenetelmiä, kuten Pellisen (2019) esittämä ja toteuttama jääkiekon laukauksen maalitodennäköisyyden mallintamisprosessi. Myös koripallon osalta löydettiin konkreettisia esimerkkejä

datan vaikutuksesta lajin evoluutioon vuosien saatossa (Abbas, 2019). Jalkapallon huipputasolta löytyi myös selostus (Arastey, 2018), jossa Brentfordin jalkapallo-organisaation toiminta mullistui data-analyysin tultua tärkeäksi osaksi joukkueen toimintaa. Valitettavasti kuitenkin esimerkiksi yksilölajien osalta havainnot jäivät tutkimuksessa hyvin vähäiseksi, mikä saattaa osittain selittyä yksilölajien lajiteknisillä seikoilla. Alussa muodostettuihin tutkimuskysymyksiin kuitenkin löydettiin vastauksia ja tutkimuksen aikana tuli käytetyistä lähteistä selkeästi esille, millaisia erilaisia tekoälyä hyödyntäviä datankeräysmenetelmiä ja teknologioita yksilöiden ja joukkueiden käytössä on, sekä miten näiden avulla kerättyä dataa hyödynnetään suorituskyvyn parantamiseksi.

Avoimesti saatavilla olevaa dataa on valtaosassa lajeista saatavilla jossain määrin ja yleensä siihen sovellettavaa yksinkertaista mallinnusprosessia on saatavilla eri joukkoviestinnän lähteistä. Esimerkiksi Kaggle, Twitter ja Medium ovat suosittuja julkaisualustoja analyysimenetelmiä harjoittaville, ja niiden kautta onkin helposti löydettävissä laji-, yksilö- sekä joukkuekohtaisia analyysejä etenkin suurimpien lajien ja liigojen osalta. Joukkueurheilun osalta yksilötason vaikutuksia ja suhtautumista tekoälyjohteiseen data-analyysiin on vaikea arvioida, sillä nämä ovat yleensä melko subjektiivisia näkemyksiä.

Uskon, että tulevaisuudessa data-analyysin merkitys kasvaa entisestään ja sitä aletaan hyödyntää yhä laajemmin myös pienemmissä sarjoissa. Datan määrä tulee todennäköisimmin kasvamaan entisestään ja keräysmetodit tulevat helpommin saataviksi. Toistaiseksi päätoimisia data-analyytikkoja on palkattu seurajoukkueisiin vain suurimmissa sarjoissa ja organisaatioissa, mutta uskoisin tämän yleistyvän jatkossa myös pienempien sarjojen ja urheiluorganisaatioiden osalta tekniikan kehittyessä. Yksilölajien osalta ongelmaksi tulevat olevat oletettavasti taloudelliset rajoitteet. Tuskin kukaan kuitenkaan silti enää kyseenalaistaa datan avulla mahdollisesti saavutettavaa etua.

Lähdeluettelo

Abbas, N. M. (2019). *NBA Data Analytics: Changing the Game - Towards Data Science*. Medium; Towards Data Science. Haettu 15.3.2023 osoitteesta https://towardsdatascience.com/nba-data-analytics-changing-the-game-a9ad59d1f116

Adidas. (2022). *adidas reveals the first FIFA World Cup™ official match ball featuring connected ball technology*. Adidas News Site | Press Resources for All Brands, Sports and Innovations. Haettu 20.2.2023, osoitteesta https://news.adidas.com/football/adidas-reveals-the-first-fifa-world-cup--official-match-ball-featuring-connected-ball-technology/s/cccb7187-a67c-4166-b57d-2b28f1d36fa0

Alamar, B. (2013). Sports analytics. New York: Columbia University Press.

Arastey, G. (2018). *Sport Performance Analysis*. Sport Performance Analysis. Haettu 14.3.2023 osoitteesta https://www.sportperformanceanalysis.com/article/2018/6/8/the-history-of-brentford-football-analytics

Aroganam, G., Manivannan, N., & Harrison, D. (2019). *Review on Wearable Technology Sensors Used in Consumer Sport Applications*. Switzerland: MDPI.

Azar, O., Morgulev, E., & Lidor, R. (2018). *Sports analytics and the big-data era*. Cham: Springer International Publishing.

Bailey, N., Bhuwalka, K., Lee, H., & Zhong, T. (2018). *Understanding Features of Successful 3 Point Shots in the NBA*. Haettu 19.3.2023 osoitteesta https://www.mit.edu/~nbailey/files/3ptreport.pdf

Barbosa, T. (2018). *Smart sport equipment: reshaping the sports landscape*. Ribeira de Pena: Desafio Singular LDA.

Barlow, J. (2015). *Data Analytics in Sports*. O'Reilly Media, Inc.

Basole, R. C., & Saupe, D. (2016). *Sports Data Visualization [Guest editors' introduction]*. IEEE Computer Graphics and Applications, *36*(5), s. 24. https://doi.org/10.1109/mcg.2016.85

Bechtold, T. (4 2021). *How the 'Idiots Who Believe' in the Analytics Movement Have Forever Changed Basketball | The Analyst*. Haettu 24.2.2023 osoitteesta The Analyst: https://theanalyst.com/eu/2021/04/how-advanced-analytics-have-changed-basketball/

Bernard, M. (2016). *Big Data in Practice (1 p.).* New York: Wiley.

Bitwise Oy. (2019). *Suomessa maailman ensimmäinen digitalisoitu jääkiekkoliiga - Liiga ja Wisehockey tekivät viiden vuoden sopimuksen - ePressi*. Haettu 26.1.2023 osoitteesta https://www.epressi.com/tiedotteet/ohjelmistoteollisuus/suomessa-maailman-ensimmainen-digitalisoitu-jaakiekkoliiga-liiga-ja-wisehockey-tekivat-viiden-vuoden-sopimuksen.html

Brefeld, U.;Davis, J.;Van Haaren, J.;& Zimmermann, A. (2019). *Machine Learning and Data Mining for Sports Analytics 5th International Workshop*, MLSA 2018, Co-located with ECML/PKDD 2018, Dublin, Ireland, September 10, 2018, Proceedings (1st ed. 2019. p.). Cham: Springer International Publishing.

Brown, S. (2021). *Machine learning, explained | MIT Sloan*. Haettu 13.2.2023 osoitteesta https://mitsloan.mit.edu/ideas-made-to-matter/machine-learning-explained

Cheema, A. (2019). *The NBA's Ongoing Three-Point Revolution - The Spax*. Haettu 17.2.2023 osoitteesta https://www.thespax.com/nba/the-nbas-ongoing-three-point-revolution/

Davenport, T. (2014). *Analytics in Sports: The New Science of Winning*.

D'Orazio, T.;& Leo, M. (2010). *A review of vision-based systems for soccer video analysis*. Pattern Recognition, 43(8), 2911–2926. https://doi.org/10.1016/j.patcog.2010.03.009

Ebling, M. (2016). *IoT: From Sports to Fashion and Everything In-Between.* IEEE Pervasive Computing, 15(4), 2. https://doi.org/10.1109/MPRV.2016.71

Euroopan parlamentti. (2020). *Mitä tekoäly on ja mihin sitä käytetään?* Haettu 19.1.2023 osoitteesta https://www.europarl.europa.eu/news/fi/headlines/society/20200827STO85804/mita-tekoaly-on-ja-mihin-sita-kaytetaan

Gain, V. (2022). *How Wimbledon is keeping its eye on the ball with IBM tech*. Silicon republic. Haettu 29.2.2023 osoitteesta *https://www.siliconrepublic.com/machines/wimbledon-2022-results-ibm-watson-technology*

Hwang, J. (2020). *Visualising evolution of the NBA offence through data*. Medium, Towards Data Science. Haettu 20.4.2023 osoitteesta https://towardsdatascience.com/visualising-evolution-of-the-nba-offence-through-data-84a30c2e3fcc

Heikkinen, S., & Nieminen, J. (2017). *Tekoäly muuttaa maailman – pian se tekee jopa lääkärin ja juristin töitä | Media- ja digitaidot | Oppiminen | yle.fi*. Haettu 19.1.2023 osoitteesta *https://yle.fi/aihe/artikkeli/2017/06/04/tekoaly-muuttaa-maailman-pian-se-tekee-jopa-laakarin-ja-juristin-toita*

Kovalchik, S. (2021). *Why Tennis Is Still Not Ready to Play Moneyball*. Issue 3.1, Winter 2021. https://doi.org/10.1162/99608f92.b665c0f4

Kohavi, R., Rothleder, N., & Simoudis, E. (2002). *Emerging trends in business analytics*. Communications of the ACM, 45(8), 45–48. https://doi.org/10.1145/545151.545177

Lewis, O. (2021). *Football: How big clubs are using data analysis to get the edge on their rivals in technology war*. Mail Online; Daily Mail. Haettu 16.3.2023 osoitteesta https://www.dailymail.co.uk/sport/football/article-9392769/Football-big-clubs-using-data-analysis-edge-rivals-technology-war.html

Linke, D., Link, D., & Lames, M. (2020*). Football-specific validity of TRACAB's optical video tracking systems*. PloS One, 15(3), e0230179–e0230179. https://doi.org/10.1371/journal.pone.0230179

Lukan, A. (2022). *Beyond the Box Score – An Intro to Hockey Analytics.* Seattle Kraken. NHL. Haettu 12.3.2023 osoitteesta https://www.nhl.com/kraken/news/beyond-box-score-intro-to-hockey-analytics/c-335471754

Macdonald, B. (2012). *An Expected Goals Model for Evaluating NHL Teams and Players*. MIT. http://hockeyanalytics.com/Research_files/NHL-Expected-Goals-Brian-Macdonald.pdf

Martin, L. (2016). *Sports Performance Measurement and Analytics: The Science of Assessing Performance, Predicting Future Outcomes, Interpreting Statistical Models, and Evaluating the Market Value of Athletes*. Hoboken: Pearson Education, Limited.

Miller, T. (2015). *Sports Analytics and Data Science: Winning the Game with Methods and Models*. Hoboken: Pearson Education, Limited.

Pellinen, J. (2019). *Maalitodennäköisyyksien mallintaminen jääkiekossa*. *http://urn.fi/URN:NBN:fi:jyu-201902251638*

Quuppa Oy. (2019a). *Wisehockey Case Study*. Haettu 15.2.2023 osoitteesta https://quuppa.com/wp-content/uploads/2019/11/Wisehockey_Case_Study.pdf

Quuppa Oy. (2019b). *Wisehockey | Quuppa - Ice Hockey Innovation - YouTube*. Haettu 15.2.2023 osoitteesta https://www.youtube.com/watch?v=uePX2WTsbHE

Runkler, T. A. (2012). *Data Analytics: Models and Algorithms for Intelligent Data Analysis (2012th ed.)*. Springer. https://doi.org/10.1007/978-3-8348-2589-6

Ryhänen, A. (2020). *Data-Analytiikan hyödyntäminen urheilujoukkueen suorituskyvyn kehittämisessä*. 22–26. https://urn.fi/URN:NBN:fi:tuni-202008186537

Schumaker, R. P., Chen, H., & Solieman, O. K. (2010*). Sports Data Mining (1. Aufl., Vol. 26)*. Springer Science + Business Media. https://doi.org/10.1007/978-1-4419-6730-5

Sellitto, C. (2015). *Big Data can give Athletes the Winning Edge*. Rockaway: Advantage Business Media.

Stanojevic, R., & Gyarmati, L. (2016). *Towards Data-Driven Football Player Assessment*. 2016 IEEE 16th International Conference on Data Mining Workshops (ICDMW), 167–172. *https://doi.org/10.1109/ICDMW.2016.0031*

Tietoevry. (2021). *Millaista dataa tarvitaan kestävyysurheilun huipulla?* Tietoevry.com. Haettu 20.3.2023 osoitteesta https://www.tietoevry.com/fi/uutishuone/kaikki-uutiset-ja-tiedotteet/artikkelit/2021/millaista-dataa-tarvitaan-kestavyysurheilun-huipulla/

Tiga, P. (2022). *Cutting-edge tech in FIFA World Cup 2022*. Haettu 20.2.2023 osoitteesta https://www.sganalytics.com/blog/cutting-edge-technologies-at-fifa-world-cup-qatar/

Tyagi, S. (2003). Using data analytics for greater profits. The Journal of Business Strategy, 24(3), 12–14. https://doi.org/10.1108/02756660310734938

Vesterinen, V., Mikkola, J., Nummela, A., Hynynen, E., & Häkkinen, K. (2009). *Fatigue in a simulated cross-country skiing sprint competition*. Journal of Sports Sciences, 27(10), 1069–1077. https://doi.org/10.1080/02640410903081860

Wisehockey Oy. (2022). *DEL implements the Wisehockey real-time sports analytics platform league-wide – DEL Managing Director: "We can create new services for our audience".* Epressi.com. Haettu 12.2.2023 osoitteesta https://www.epressi.com/tiedotteet/tietotekniikka/del-implements-the-wisehockey-real-time-sports-analytics-platform-league-wide-del-managing-director-we-can-create-new-services-for-our-audience.html

Xu, S., & Lee, S. (2022). An Inertial Sensing-Based Approach to Swimming Pose Recognition and Data Analysis. *Journal of Sensors*, *2022*, 1–12. https://doi.org/10.1155/2022/5151105