

Teemu Korhonen

**Tiedonlouhinnan hyödyntäminen digitaalisessa
markkinoinnissa**

Tietotekniikan Kandidaatintutkielma

28. huhtikuuta 2021

Jyväskylän yliopisto

Informaatioteknologian tiedekunta

Tekijä: Teemu Korhonen

Yhteystiedot: temikorh@student.jyu.fi

Ohjaaja: Tuomo Rossi

Työn nimi: Tiedonlouhinnan hyödyntäminen digitaalisessa markkinoinnissa

Title in English: Utilizing data mining in digital marketing

Työ: Kandidaatintutkielma

Opintosuunta: Tietotekniikka

Sivumäärä: 21+0

Tiivistelmä: Tässä kandidaatintutkielmassa tarkastellaan kuinka tiedonlouhintaa voidaan hyödyntää digitaalisessa markkinoinnissa ja kuinka tiedonlouhinta-algoritmilli loughittua dataa voidaan hyödyntää asiakkaiden profiloinnissa. Tutkielmassa selvitetään miten assosiaatiosääntöjenlouhintamenetelmiin pohjautuva markkinakoriantalyysi ilmenee digitalisen markkinoinnin asiakasprofilointiprosessissa ja kuinka markkinoijat voivat sitä hyödyntää osana tätä prosessia. Lisäksi pyritään selvittämään käytännön esimerkkien kautta tapoja hyödyntää tai käyttää apuna tätä kyseistä tiedonlouhintamenetelmää kun asiakkaille personoidaan digitaalista markkinointia.

Avainsanat: Digitaalinen markkinointi, tiedonlouhinta, assosiaatiosääntö, markkinakoriantalyysi

Abstract: This bachelor's thesis examines how data mining can be utilized in digital marketing and how data mined with the data mining algorithm can be utilized in customer profiling. The thesis explores how market basket analysis based on association rules mining methods manifests itself in digital marketing customer profiling process and how marketers can utilize it as part of this process. In addition, the aim is to find out, through practical examples, how this particular data mining method can be utilized or used as an aid when personalizing digital marketing to customers.

Keywords: Digital marketing, data mining, association rule, market basket analysis

Kuviot

Kuvio 1. Profilointiprosessi (Merisavo 2006)	4
Kuvio 2. Yksinkertainen tapahtumätietokanta malli (Kantardzic 2011)	7
Kuvio 3. Digitaalisen markkinoinnin aihealueita, jossa hyödynnetään datatieteitä ja tiedonlouhintaa (Saura 2020)	8
Kuvio 4. Markkinakorianalyysi (Tuffery 2011)	11

Sisällys

1	JOHDANTO	1
2	DIGITAALINEN MARKKINOINTI	2
	2.1 Digitaalisen markkinoinnin käsitteet ja hyödyt	2
	2.2 Asiakkaiden profilointi digitaalisessa markkinoinnissa	3
3	TIEDONLOUHINTA.....	5
	3.1 Tiedonlouhinnan menetelmät	5
	3.2 Tiedonlouhinta liiketoiminnassa ja digitaalisessa markkinoinnissa.....	7
4	MARKKINAKORIANALYYSIN HYÖDYNTÄMINEN ASIAKKAIDEN PRO- FILOINNISSA.....	10
	4.1 Markkinakorianalyysi osana profilointiprosessia	10
	4.2 Markkinakorianalyysin käytännönsovellutuksia	13
5	YHTEENVETO.....	14
	LÄHTEET	16

1 Johdanto

Digitaalinen markkinointi on parin viime vuoden aikana nostanut päätään ja on yksi nopeiten kasvavista tietotekniikkaan nojaavista aihe-alueista. Sen osuus koko medianostuksesta on tällä hetkellä 36,6 prosenttia. Samaan aikaan digitalisaation ja tietoyhteiskunnan kautta maailmassa olevan datan määrä on kasvanut ja kasvaa vieläkin. Tämä tarjoaa markkinoijille erinomaiset lähtökohdat kerätä tietoa nykyisistä ja potentiaalisista asiakkaista. Tästä syystä tutkielman aiheeksi valikoitui tiedonlouhinnan hyödyntäminen digitaalisessa markkinoinnissa.

Tiedonlouhinnan menetelmien hyödyntämisellä osana digitaalista markkinointia voitaisiin säästää markkinoijien ja yritysten aikaa ja resursseja, koska läpikäytävän tiedon määrä on valtava ja kilpailu kovaa digitaalisilla markkinoilla. Digitaalisen markkinoinnin ajankohtaisuuden puolesta puhuu ennen kaikkea koronapandemian vauhdittama verkkokauppojen kasvu, joka kasvatti Suuri verkkokauppatutkimus 2020 - tutkimuksen mukaan verkko-ostamista jo pelkästään suomalaisten keskuudessa 30 prosentilla (Posti 2020, s.8). Paytrailin (2021, s.5) tuottaman Verkkokauppa Suomessa 2020 - tutkimuksen pohjalta Suomen verkkokaupan kokonaisliikevaihdon arvioidaan olleen vuonna 2020 11,8 miljardia euroa, joka korostaa verkossa asioinnin laajuutta ja sitä kautta digitaalisen markkinoinnin tärkeyttä.

Tämä tutkielma pyrkii siis selvittämään mitä on digitaalinen markkinointi ja tiedonlouhinta, ja kuinka tiedonlouhinta voidaan hyödyntää digitaalisen markkinoinnin asiakkaiden profiloinnissa. Tiedonlouhinnan hyödynnettävyyttä tässä asiassa tarkastellaan markkinakorianalyysimenetelmän pohjalta, joka pohjautuu assosiaatiosääntöjenlouhintamenetelmiin. Aihetta tarkastellaan tutkielman laajuuden vuoksi vain hyötynäkökulmasta ja tiedonlouhinta-menetelmistä tarkempaan tarkasteluun otetaan vain assosiaatiosäännöt. Tutkielmassa pyritään vastaamaan tutkimuskysymykseen, kuinka hyödyntää markkinakorianalyysia asiakkaiden profiloinnissa. Tähän pyritään vastaamaan tutkimusstrategian eli systemaattisen kirjallisuuskatsauksen avulla taustakirjallisuuteen pohjaten ja liittäen markkinakorianalyysin ja sen assosiaatiosäännöt käytännön esimerkein ymmärrettävämpään muotoon.

2 Digitaalinen markkinointi

Nopeasti digitalisoituvassa maailmassa yhä useampia asiaa linkittyy jollain lailla tietotekniikkaan. Digitalisaatio on johtanut mm. markkinointimaiseman muutokseen. Markkinointi on jo pitkään ollut osa yritysten DNA:ta ja nykyään sen yhdeksi osa-alueeksi tunnustetaan digitaalinen markkinointi (eng. digital marketing), jossa internetillä on keskeinen rooli tietojen integroinnissa toimittajien, asiakkaiden ja organisaation välillä. Tietokoneiden, matkapuhelinten ja muiden tietoteknisten laitteiden johdosta kuluttajat ovat siirtyneet entistä enemmän kivijalkaliikkeissä asiointista kohti verkkokauppa-asiointia. Suomalaisista lähes 60 prosenttia ostaa verkosta kuukausittain, jo lähes kolmannes puolestaan viikoittain (Posti 2020, s.7). Päivittäistavarakaupan puolella on viime vuoden aikana huomattu huomattavaa verkkokaupan osuuden kasvua, kun sen osuus oli kolminkertaistunut vuoden 2019 lukuun nähden (Päivittäistavarakauppa ry 2021). Kuluttajien siirtyessä verkkoon tulee myös markkinoijien siirtyä mukana ja integroida vanhat markkinointitavat uusiin, jotka pätevät ja toimivat erilaisissa digitaalisten medioiden alustoilla. Tähän kaikkeen digitaalinen markkinointi pyrkii vastaamaan. Tässä luvussa pureudutaankin siihen mitä digitaalinen markkinointi on ja kuinka asiakkaiden profilointi ja sitä kautta personoidun markkinointiviestinnän jakaminen siinä tapahtuu.

2.1 Digitaalisen markkinoinnin käsitteet ja hyödyt

Digitaalisella markkinoinnilla tarkoitetaan uusien markkinointiviestinnän digitaalisten muotojen ja medioiden, kuten internetin, mobiilimedian ja muiden interaktiivisten kanavien hyödyntämistä osana markkinointiviestintää (Karjaluo 2010, s.13-14). Digitaalinen markkinointi on siis kaikkea digitaalisilla media-alustoilla tapahtuvaa markkinointia. Juuri markkinoinnin ulottuminen kaikkiin digitaalisen median alustoihin ja kanavoihin erottaa sen verkkomainonnasta, jossa mainostetaan ja viestitään pelkästään internet kanavien kautta.

Digitaalinen markkinointi käsittää useita muotoja, joista tunnetuimmat ovat sähköinen suoratoistomarkkinointi eli sähköposti ja mobiili: lähinnä tekstiviestit ja internetmarkkinointi, joka kattaa yrityksen verkkosivut, erilaiset kampanjasivustot, verkkomainonnan kuten ban-

nerit, ja hakukonemarkkinoinnin (Karjaluoto 2010, s.14). Yrityksen omat verkkosivut, mobiilimainonta ja hakukoneoptimointi (SEO) ovat erittäin tärkeitä kanavia ottaa huomioon yrityksen digitaalisessa markkinoinnissa. Digitaalinen teknologia mahdollistaa myös markkinoinnissa monien sähköisten kanavien käytön samanaikaisesti ja helpottaa yrityksen tapaa markkinoida profiloituneesti ja keskitetysti (Isohookana 2007, s.251-252).

Digitaalinen markkinointi on ennen kaikkea muuttanut sitä, kuinka yritykset nykypäivänä kommunikoivat sidosryhmilleen. Digitaalisten alustojen sisällyttäminen yhä enemmän yritysten markkinointisuunnitelmiin ja jokapäiväiseen toimintaan sekä asiakkaiden alati kasvava digitaalisten laitteiden käyttäminen ja verkkovierailut fyysisten kauppojen vierailun sijaan ovat yleistäneet digitaalisten markkinointikampanjoiden käyttämistä. Markkinointia voidaan myös selkeästi personoidummin kohdentaa tietyille asiakkaille.

Markkinoijat hyötyvät digitaalisesta markkinoinnista esim. kustannuksissa, koska viestintäkustannukset voivat olla huomattavasti pienemmät sähköpostin tai mobiiliviestinnän kuin perinteisen kirjepostin kautta. Muita selkeästi havaittavia hyötyjä ovat markkinoinnin vuorovaikutus, globaalit markkinat verkossa, kohdistettu markkinointi, asiakkuuksien hallinta ja se ettei markkinointi ei ole yhtä helposti kilpailijoiden nähtävissä. Kuluttajat myös hyötyvät digitaalisesta markkinoinnista, koska se helpottaa tiedon etsintää, mahdollistaa ostosten tekemisen paikasta sekä ajasta riippumatta ja tarjoaa kohdistettua sekä personoitua eli kuluttajien henkilökohtaisten preferenssien mukaista markkinointiviestintää.

2.2 Asiakkaiden profilointi digitaalisessa markkinoinnissa

Profilointiprosessi digitaalisessa markkinointiviestinnässä koostuu vuorovaikutuksesta asiakkaan kanssa, asiakastietojen hankkimisesta, joka voi olla peräisin esimerkiksi yrityksen asiakasrekisteristä, tietojen analysoinnista, markkinointimixin¹ profiloinnista asiakkaan profiilin perusteella sekä markkinointitoimien kohdistamisesta (Merisavo 2006, s.105-106). Asiakastietoja voidaan kerätä asiakkaasta erilaisten vuorovaikutustilanteiden kuten ostotapahtuman

1. Markkinointimix on markkinointitoimintaa, jota käytetään luomaan, viestimään ja tuomaan lisäarvoa asiakkaalle. Alun perin se koostui neljästä markkinointimix-muuttujasta eli 4P:stä, jotka laajennettiin myöhemmin seitsemään. (Kotler ym. 2010, s.866)

tai asiakkaan yrityksen verkkosivuilla toteuttaman käyttäytymisen pohjalta. Asiakastietoja saadaan myös tietokannoista mm. erilaisten tiedonlouhintamenetelmien pohjalta (Chaffey 2019, s.247). Tällainen digitaalisen markkinoinnin personoinnin kannalta olennainen asiakkaiden profilointiprosessi esitellään Kuviossa 1.



Kuvio 1. Profilointiprosessi Merisavon (2006) esimerkin mukaisesti esitettynä.

Esimerkiksi erilaiset digitaaliset markkinointiviestit kuten sähköpostit ja mobiiliviestit on mahdollista personoida asiakkaan mielenkiinnon ja kiinnostuksen kohteiden mukaisesti, jonka seurauksena asiakas saa tietoa vain niistä tuotteista ja palveluista, joista hän haluaa tietää. Juuri tällaisen henkilökohtaisen personoinnin avulla voidaan huomattavasti vahvistaa markkinoinnin toimivuutta. (Merisavo 2006, s.55) Asiakkaiden profilointiprosessin jäljiltä markkinoijat pystyvät personoimaan asiakaskohtaisesti millaista digitaalista markkinointia ja markkinointiviestintää heille kohdistetaan. Profiloinnista on myös hyötyä myynninedistämässä, jossa tiettyjä palveluita ja tuotteita voidaan kohdistaa asiakasprofiilin mukaisesti vain tietyille asiakasryhmälle.

3 Tiedonlouhinta

Tiedon ja saatavilla olevan datan määrä on kasvanut vuosi vuodelta digitalisaation myötä eikä kasvulle näy loppua. Koska tällaisten suurien tieto- ja datamäärien läpikäyminen tai kirjaaminen ei ole mielekästä fyysisesti, on sitä varten kehitetty erilaisia tiedonlouhintamenetelmiä (eng. data mining) helpottamaan tiedon läpikäymistä. Tiedonlouhinnallekin on olemassa useita erilaisia määritelmiä, mutta usein tiedonlouhinnan määritellään olevan liiketoimintaprosessi, jolla tutkitaan suuria tietomääriä mielekkäiden mallien ja sääntöjen löytämiseksi (Linoff ja Berry 2011, s.2).

Tiedonlouhinnan tarkoituksena on siis löytää suuresta määrästä dataa oleellinen tieto erilaisien mallien ja sääntöjen avulla, ja muokata se sitten yleistajuisempaan muotoon. Tiedonlouhinta kuvaillaan ihmisten ja tietokoneen yhteistulokseksi, ja parhaat tulokset saavutetaankin tasapainottamalla inhimillisten asiantuntijoiden tietämys kuvailemalla ongelmia ja tavoitteita tietokoneiden hakutoiminnoilla (Kantardzic 2011, s.2).

Tiedonlouhinta voi olla manuaalista tai automaattista, jolloin siinä on taustalla koneoppiminen. Tiedonlouhinnassa tiedot tallennetaan elektronisesti ja haku on automatisoitu tai ainakin täydennetty tietokoneella (Witten, Frank ja Hall 2011, s.4). Tänä päivänä tiedonlouhintatekniikat pitkälti pohjautuvat koneoppimisen ja tekoälyn puolella kehitettyihin algoritmeihin.

3.1 Tiedonlouhinnan menetelmät

Tiedonlouhinnan suurimmat hyödyt tulevat siis esiin suurien tietomääriä läpikäydessä, mutta kuinka tämä oikein tapahtuu ja millaisia ovat tiedonlouhintamenetelmät? Mallin löytämisprosessi, kun louhitaan tietojoukkoja, tunnetaan nimellä 'analyttinen mallinnus', tiedonlouhintamallin luomiseksi (Leventhal 2010, s.139). Prosessi pitää sisällään suhteiden tunnistamista datan muuttujien väliltä ja näitä suhteita käytetään ennustavien tai kuvaavien mallien luomiseen. Tulos ilmaistaan kaavana tai algoritmina, joka voi laskea pisteet (ennustettu arvo tai todennäköisyys) jokaiselle yksittäiselle tietueelle, esimerkiksi vasteelle, virheelle tai toistuvalla myynnillä, kyseisen tietueen data-arvojen mukaan (Leventhal 2010, s.139).

Tiedonlouhintamallit jaetaan yleensä kahteen päämalliin niiden päätavoitteiden mukaisesti, jotka ovat ennustava ja kuvaava malli. Leventhal (2010, s.140) mukaan ennustava malli on tiedonlouhintamalli, joka on suunniteltu ennustamaan tietty tulos tai kohdemuuttuja. Kuvaava malli puolestaan on malli, joka antaa paremman käsityksen tiedoista ilman yksittäistä kohdemuuttujaa. Ennustavan mallin mukaisia tiedonlouhintatekniikoita ja menetelmiä ovat mm. regressiot ja päätöspuumallit, kun taas kuvaavan mallin mukaisia menetelmiä ja tekniikoita ovat esim. klusterointi ja assosiaatiosäännöt. Tyypillisimpiä digitaalisessa markkinoinnissa puolestaan käytettäviä tiedonlouhintatekniikoita ja menetelmiä ovat klusterointi, assosiaatiosäännöt ja luokittelu. Näistä tarkemmin perehdytään assosiaatiosääntöjen menetelmiin.

Assosiaatiosäännöt (eng. association rules) edustavat joukkoa suhteellisen uusia menetelmiä, jotka sisältävät algoritmeja, kuten markkinakorianalyysin, apriori-algoritmin ja WWW-polku-kulkumalleja (Kantardzic 2011, s.105). Tarkoituksena assosiaatiosääntöihin pohjautuvissa menetelmissä on löytää mielenkiintoisia suhteita muuttujien väliltä halutuista yleensä todella suurista tietokannoista.

Agrawal, Imieliński ja Swami (1993, s.208) mukaan assosiaatiosääntöjen louhinta voidaan määritellä muodollisesti seuraavasti: $I = \{i_1, i_2, \dots, i_n\}$ on joukko n binäärisiä määritteitä, joita kutsutaan kohteiksi tai esineiksi (eng. items). $D = \{t_1, t_2, \dots, t_m\}$ on joukko tapahtumia, joita kutsutaan tietokannaksi (eng. database). Jokaisella tapahtumalla D :ssä on ainutlaatuinen tapahtumatunnus eli TID ja sisältää osan I osioista. Sääntö sitten implikoidaan matemaattiseen muotoon $X \Rightarrow Y$, jossa $X, Y \subseteq I$ ja $X \cap Y = \emptyset$ eli jos tapahtuu tapahtumat X ja Y , tapahtuma I tapahtuu jollain todennäköisyydellä. Kuviossa 2 on nähtävillä miltä tällainen yksinkertainen tapahtumatietokanta malli näyttää.

Han (2012, s.247) mukaan assosiaatiosääntöjen louhinta voidaan pitää kaksivaiheisena prosessina. Ensimmäiseksi täytyy löytää kaikki usein käytettävät osajoukot: Määritelmän avulla kutakin näistä joukoista esiintyy ainakin yhtä usein kuin ennalta määrätty vähimmäistuki (min_sup). Toisessa vaiheessa luodaan vahvat assosiointisäännöt usein olevista kohteista: Määritelmän mukaan näiden sääntöjen on täytettävä vähimmäistuki ja vähimmäisluottamus (min_conf). Sääntöjä, jotka täyttävät sekä vähimmäistukikynnyksen (min_sup) että vähimmäisluottamuskynnyksen (min_conf), kutsutaan voimakkaiksi. Tyypillisesti assosiaatiosääntöjä pidetään mielenkiintoisina, jos ne täyttävät sekä vähimmäistukikynnyksen että

Database DB:

TID	Items
001	A C D
002	B C E
003	A B C E
004	B E

Kuvio 2. Yksinkertainen tapahtumatietokantamalli Kantardzicin (2011) mallin mukaisesti esitettynä.

vähimmäisluottamuskyngyksen (Han 2012, s.254). Ne myös puolestaan vastaavat löydettyjen sääntöjen hyödyllisyyttä ja varmuutta. Tavanomaisesti tuki- ja luottamusarvot merkitään siten, että ne esitellään välillä 0% - 100% ja joissain tapauksissa välillä 0-1,0. Käyttäjän tai järjestelmän ylläpitäjän on mahdollista määrittellä nämä raja-arvot itse.

Han (2012, s.246) esimerkin mukaisesti mm. tieto siitä, että tietokoneita ostavat asiakkaat ostavat yleensä myös virustorjuntaohjelmistoja samanaikaisesti, esitetään seuraavassa assosiaatiosäännössä: tietokone \implies virustentorjuntaohjelmisto [tuki = 2%, luottamus = 60%]. 2 prosentin tuki säännölle tarkoittaa, että 2 prosenttia kaikista analysoiduista tapahtumista osoittaa, että tietokone- ja virustorjuntaohjelmistot ostetaan yhdessä. 60 prosentin luottamus tarkoittaa, että 60 prosenttia tietokoneen ostaneista asiakkaista osti myös ohjelmiston.

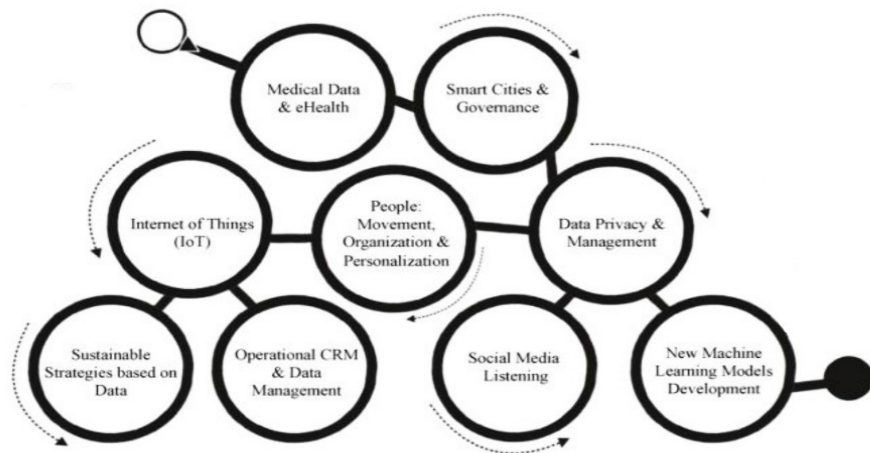
3.2 Tiedonlouhinta liiketoiminnassa ja digitaalisessa markkinoinnissa

Tiedonlouhinta käsittelee useasti dataa, joka on kerätty johonkin muuhun tarkoitukseen kuin itse louhintaan ja sillä kerätty aineisto pyrkii vastaamaan tiettyyn kysymykseen. Usein käyttökohteina ovat asiakkuuksienhallinta, ihmisten käyttäytymisen tai markkinoinnin analysointi. Kuten edellisessä luvussa jo kävikin ilmi, tiedonlouhintamenetelmiä ja tekniikoita on monia erilaisia ja toiset sopivat paremmin toisiin tilanteisiin kuin toiset. Onnistuneita menetelmiä voidaan hyödyntää monilla eri alueilla, kuten personoiduista kampanjoista, myyntikäytännöistä, uusista tuotteista, risti- ja lisämyynneistä sekä markkinajärjestelyistä tiedon-

louhinnalla saatujen tietojen markkinoinnissa (Tekin ym. 2019, s.50). Onkin tärkeää valita oikeanlaiset menetelmät ja tekniikat, jotka sopivat yritykselle tai kyseiseen aiheeseen.

Tiedonlouhinnan käytännöllinen seuraus on, että organisaatiot, jotka haluavat menestyä tietojensa sekä datansa hyödyntämisessä liiketoimintansa parantamisessa, eivät pidä tiedonlouhintaa pelkkänä sivutuotteena. Sen sijaan heidän liiketoimintastrategiansa on sisällytettävä tietojen kerääminen, tietojen analysointi pitkäaikaista hyötyä varten ja toimiminen tulosten perusteella (Linoff ja Berry 2011, s.3). Tiedonlouhinnan hyödyntämisen pitäisikin siis olla osa jokaisen datasta kiinnostuneen yrityksen liiketoimintastrategiaa. Näin se olisi myös helppo ottaa osaksi digitaalisen markkinoinnin menetelmiä.

Kuviossa 3 on esitelty muita aihealueita digitaalisen markkinoinnin saralta, jossa on hyödynnetty datatieteitä ja tiedonlouhinnan menetelmiä. Siitä saakin hyvän käsityksen kuinka monien eri aiheiden markkinoinnissa erilaisten tiedonlouhinnan menetelmien hyödyntäminen onkaan mahdollista ja kuinka laajalti sen menetelmät ovat hyödynnettävissä.



Kuvio 3. Digitaalisen markkinoinnin aihealueita, joissa hyödynnetään datatieteitä ja tiedonlouhintaa Sauran (2020) mukaisesti kuvattuna.

Digitaalisen markkinoinnin saralla yritysten omat asiakastietokannat ovat esimerkiksi erin-

omaisia tiedon ja datan lähteitä ja monen yrityksen tulisikin entistä enemmän panostaa sieltä saatavan tiedon hyödyntämiseen. Asiakasprofiilien ja mahdollisen personoidun markkinointiviestinnän suuntaaminen asiakkaille on ehkäpä potentiaalisimpia ja tehokkaimpia tiedonlouhinnan hyödyntämiskohteita digitaalisessa markkinoinnissa. Erilaiset markkinakorianaalyytit ovat olleet pitkään markkinoijien eniten suosima tiedonlouhinnan algoritmeja hyödyntävä menetelmä.

4 Markkinakorianalyysin hyödyntäminen asiakkaiden profiloinnissa

Jo aiemmin mainittu markkinakorianalyysi (eng. market basket analysis), joka on assosiaatiösääntöjen pohjalta tehty menetelmä, on yksi eniten markkinoinin piirissä jo pidemmän aikaa käytettyjä tiedonlouhintamenetelmiä. Tässä tapauksessa markkinakori, joka on hyvin määritelty liiketoimintatapahtuma on kokoelma tuotteita, jotka asiakas ostaa yhden asiakastapahtuman aikana (Kantardzic 2011, s.282). Digitaalisen markkinoinnin kontekstissa asiakkaan vierailu esimerkiksi yrityksen verkkosivuilla tai verkkokaupassa on tällainen asiakastapahtuma.

4.1 Markkinakorianalyysi osana profilointiprosessia

Markkinakorianalyysissa assosiaatiösääntöjen avustuksella koetetaan saada selvyys keitä asiakkaat ovat heidän ostamiensa ja käyttämiensä hyödykkeiden perusteella, ja tunnistamaan asiakkaita, jotka todennäköisesti ovat valmiita ostamaan muitakin hyödykkeitä ja palveluita. Pyritään siis profiloimaan asiakkaita heidän ostostensa mukaan ja löytämään suhteita eri hyödykkeiden väliltä. Profilointi tarkoittaa tässä kontekstissa käyttäjien tietojen keräämistä talteen kauppapaikassa säilytettävään asiakasprofiiliin ja sopivan tuotteen tarjoamista tähän profiiliin perustuen. Toisin sanoen käyttäjän erilaisista ominaisuuksista ja/tai mielenkiinnon kohteista muodostetaan kuva (eli profiili) hänestä kerättyjen tietojen pohjalta, käyttäen apuna mahdollisesti myös muiden käyttäjien tietoja (Laveri 2006, s.12).

Markkinakorianalyysissa muuttujat ovat tuotteiden indikaattoreita, ja sääntöjä sovelletaan indikaattoreihin, jotka ovat yhtä suuria kuin 1, toisin sanoen ostettuihin tuotteisiin (Tuffery 2011, s.287). Jos tiedetään, että tuotteen X ostajat ostavat hyvin todennäköisesti tuotetta Y X:än ostamisen yhteydessä ja jos asiakas ostaa tuotetta X mutta ei tuotetta Y, sanomme, että hän on potentiaali Y-asiakas. Tämä tekniikka auttaa ymmärtämään suhdetta eri kohteiden eli tuotteiden välillä, joita asiakas sijoittaa ostoskoriin ja asiakkaiden ostokäyttäytymiseen (Tekin ym. 2019, s.56). Yritykset voivat käyttää näiden suhteiden ja liittymien tuntemista myöhemmin voittojen kasvattamiseen. Erilaisia tiedonlouhinnalla saatuja ennusteita on

myös hyödyllistä käyttää liiketoiminnan kehittämisessä.

	Product 1			Product 2			...
Customer A	Type	Brand	Purchase date	Type	Brand	Purchase date	...
Customer B	Type	Brand	Purchase date	Type	Brand	Purchase date	...
Customer C	Type	Brand	Purchase date	Type	Brand	Purchase date	...
...
...

Kuvio 4. Markkinakoriantalyysi Tufferyn (2011) esimerkin mukaisesti kuvattuna.

Kuvio 4 on yksinkertainen esimerkki markkinakoriantalyysistä (Tuffery 2011). Siellä voisi olla useampikin muuttuja kuin nyt esitelty ja tuotteita olisi useimmiten huomattavasti enemmän koska tietokannat, joita louhitaan olisivat todella suuria. Product X voisi olla vaikka kitara ja product Y kitarankielet josta voitaisiin johtaa assosiaatiosääntö: kitara \implies kitarankielet [tuki = 15%, luottamus = 75%]. Tämä siis kertoisi meille, että 15% kaikista analysoiduista tapahtumista osoittaa, että kitara ja kitarankielet ostetaan yhdessä. 75 prosentin luottamus tarkoittaa, että 75 prosenttia kitaran ostaneista asiakkaista osti myös kielet. Näiden kahden tuotteen väliltä löytyy vahva assosiaatio, jota voidaan hyödyntää vuorovaikutuksessa asiakkaan kanssa osana profilointiprosessia.

Aiemmin esitelty Merisavon (2006) profilointiprosessin kaaviota tarkastelemalla voidaan selkeästi löytää yhteys markkinakoriantalyysin hyödyntämiseen. Kaikki profilointiprosessissa lähtee asiakkaasta, joka alkaa vuorovaikuttamaan esim. asioimalla yrityksen verkkosivulla. Asiakas saattaa kiinnostua erilaisista tuotteista ja sijoittaa niitä suosikkilistalle tai ostoskoriin. Kaikella tällä toiminnallaan asiakas tarjoaa yritykselle ja sen markkinoijille dataa, jota louhia asiakasprofilointia varten.

Markkinakoriantalyysin käyttäminen tulee esille vuorovaikutusvaiheessa. Tässä vaiheessa

pyritään yllä mainitun kitaraesimerkin mukaisesti louhia tuotteita sisältävästä tietokannasta sääntöjä, jotka täyttäisivät käyttäjien ja järjestelmien ylläpitäjien määrittämän vähimmäistuen ja vähimmäisluottamuksen raja-arvot ja sitä kautta täyttäisivät hyvän assosiaatiosääntön vaatimukset. Tätä kautta selvitettyt assosiaatiot ja suhteet tuotteiden tai hyödykkeiden välillä auttavat markkinoijia kohdistamaan seuraavassa vaiheessa markkinointiviestintää kyseisiä hyödykkeitä ostaneille asiakkaille ja profiloimaan heitä ostotensa tai samankaltaisia ostoksia tehneiden asiakkaiden avulla.

Yleensä tämän jälkeen profiointiprosessissa prosessoidaan kaikki saatavilla oleva tieto yhdeksi asiakasprofiiliksi, jonka jälkeen kustomoidaan asiakasprofiilin perusteella haluttua markkinointiviestintää. Kustomointi tarkoittaa tässä kontekstissa tarkasti harkitun ja kiinnostavan digitaalisen markkinointisisällön tekemistä asiakkaalle eli markkinointitulosten tuotantoa. Markkinointituloste on useimmiten digitaalisissa kanavissa mm. personoitu sähköposti, mobiiliviesti tai kustomoitu verkkosivu. Viimeisenä tapahtuu toimitus, jossa asiakkaalle saatetaan markkinakorianalyysin pohjalta saatujen tietojen perusteella personoitu viesti esim. potentiaalisiin tuotteisiin Y liittyen tai muuta vastaavaa.

Markkinakorianalyysista saaduilla tiedoilla pystytään siis profiloimaan asiakas hänen edellisten tai muiden asiakkaiden ostotapahtumien ja ostamien hyödykkeiden pohjalta ja jatkossa personoidun markkinointiviestinnän kautta tarjota kyseisiin hyödykkeisiin liittyviä tarjouksia ja muuta viestintää. Tulevista ostoista pystytään sitten saamaan lisää tietoa ja parantamaan profiointia entisestään.

Tekin ym. (2019, s.56) mukaan tämänkaltaisilla tiedonlouhintamenetelmillä ja analyyseillä voidaan myös saavuttaa myyntiennusteet seuraavia kuukausia varten jokaiselle tuotteelle, tuotteelle voidaan tehdä markkinointikampanjoita ja yhdessä ostetuille tuotteille voidaan tehdä myyntikokonaisuuksia. Asiakkaat voidaan ryhmitellä ostamiensa tuotteiden mukaan, ja uuden tuotteen potentiaaliset asiakkaat voidaan tunnistaa sekä profiloida.

Assosiaatiosääntömenetelmät ja markkinakorianalyysi tarjoavat siis laajalti erilaisia hyödyn-tämismahdollisuuksia asiakkaiden profiointiin ja heille personoidun markkinointiviestinnän ulkopuolellekin ennen kaikkea myynninedistämistoimenpiteisiin liittyen. Tämän pohjalta heille voidaan suorittaa profiointiprosessi, jonka kautta saadaan selville millaista personoi-

tua digitaalista markkinointia ja markkinointiviestintää heille voitaisiin kohdistaa.

4.2 Markkinakorianalyysin käytännönsovellutuksia

Käytännössä tällaista markkinakorianalyysia hyödynnetään esimerkiksi verkkokauppojen kuten Amazonin asiakastilin ominaisuutena luetteloida mahdollisesti kiinnostavia tuotteita. Asiakas sitten saa ilmoituksen kun muut asiakkaat, jotka ostivat hänen ostamansa tuotteen, tarkastivat tai ostivat myös toisen luettelon tuotteista. Asiakkaan mielenkiintoa ja ostohalukkuutta pyritään näin herättelemään suorittamaan mahdollisia lisäostoja. Markkinakorianalyysin avulla kerätyn tiedon pohjalta tai tässä tapauksessa kiinnostavien tuotteiden listan pohjalta luodun asiakasprofiilin avulla voidaan asiakkaalle tarjota vain häntä kiinnostavia tarjouksia, tuotteita ja tarjota personoitua markkinointiviestintää.

Markkinakorianalyysia voidaan hyödyntää tätä kautta erittäin tehokkaasti kohdistamaan markkinointikampanjoita tietyille asiakkaille. Asiakkaiden omien tai samankaltaisten asiakasprofiilien perusteella heidät voidaan houkutella myös ostamaan vastaavia tuotteita äskettäin ostamilleen tuotteille. Lisäksi verkkosivuja voidaan kustomoida asiakaskohtaisesti mahdollisten tuotteiden väliltä löytyvien assosiaatioiden ja heidän ostohistoriansa perusteella.

5 Yhteenveto

Tutkielman tarkoituksena oli tarkastella ja esitellä taustakirjallisuuteen perustuen digitaalisen markkinoinnin tunnuspiirteitä ja asiakkaiden profilointia siinä, sekä tiedonlouhinnan menetelmiä ja hyödyntämistä osana digitaalista markkinointia. Aiheen laajuuden vuoksi tarkasteluun valikoitui tiedonlouhinnan menetelmistä assosiaatiosäännöt ja kuinka niiden pohjalta toteutettavaa markkinakorianalyysia voidaan hyödyntää osana asiakkaiden profilointiprosessia. Tutkielmassa haluttiin demonstroida käytännön esimerkin kautta kuinka assosiaatiosääntöjä hyödynnetään digitaalisen markkinoinnin asiakasprofiloinnin ja personoidun viestinnän toteuttamisessa. Tutkielmassa päädyttiin tarkastelemaan aihetta vain hyötynäkökulmasta tutkielman suppeuden vuoksi, joten haitat ja heikkoudet jätettiin tarkoituksella tarkastelematta.

Taustakirjallisuuteen pohjaten pystyttiin havaitsemaan, että tiedonlouhinnan menetelmät ovat hyödyllisiä käyttää osana digitaalista markkinointia. Varsinkin asiakkaiden profilointiin ja heille personoidun markkinointiviestinnän tekemiseen on markkinakorianalyysi erittäin hyvä menetelmä hyödyntää, kuten luvussa 4 huomattiin. Tiedonlouhinnan hyödyntäminen osana digitaalista markkinointia havaittiin myös helpottavan markkinoiden tehtäviä, koska se automatisoi tiedon keruuta. Sen avulla on myös saatavissa yrityksen näkökulmasta muutakin kuin asiakkaiden profilointia hyödyttävää tietoa.

Tutkimusstrategiana kirjallisuuskartoitus toimii hyvin, mutta tiedonlouhinta ja digitaalinen markkinointi aiheena vaatisi myös konkreettisia testejä sekä toteutuksia luotettavien ja entistä hyödyllisempien tulosten pohjaksi. Alati kasvavan tiedon määrän vuoksi olisi myös yritysten tarpeellista perehtyä ja tutkia tiedonlouhinnan hyödyntämistä vielä tarkemmin osana digitaalista markkinointia.

Tiedonlouhinnan toteuttamiseen liittyy tietenkin paljon eettisiä kysymyksiä, mutta niihin ei pystytty tämän tutkielman laajuudessa pureutumaan. Eettiset kysymykset kuitenkin tarjoavat ajankohtaisen ja laajan aiheen jollekin asiasta kiinnostuneelle perehtyä omassa tutkimuksessaan. Näistä syistä jatkotutkimuskohteena voisi toteuttaa markkinakorianalyysin vaikkapa jollekin digitaalisessa mediassa operoivalle verkkokaupparyitykselle, jonka avulla voidaan tarkastella tuotteiden assosiaatioita ja päivittää asiakkaiden asiakasprofileja ja koh-

dentaa heille tarkoitettua markkinointiviestintää. Tiedonlouhinnan menetelmien hyödynnettävyyden tutkiminen digitaalisen markkinoinnin osana ja apuna tarjoaa vielä paljon aiheita syventyä jatkossa.

Lähteet

Agrawal, Rakesh, Tomasz Imieliński ja Arun Swami. 1993. "Mining Association Rules between Sets of Items in Large Databases". *SIGMOD Rec.* (New York, NY, USA) 22, numero 2 (kesäkuu): 207–216. ISSN: 0163-5808. <https://doi.org/10.1145/170036.170072>.
<https://doi.org/10.1145/170036.170072>.

Chaffey, Dave. 2019. *Digital marketing : strategy, implementation and practice*. Seventh edition. Ebook Central SUPO/MUPO. Harlow, England: Pearson. <https://ebookcentral.proquest.com/lib/jyvaskyla-ebooks/detail.action?docID=5672727>.

Han, Jiawei. 2012. *Data mining : concepts and techniques*. 3rd ed. Toimittanut Jian Pei. Morgan Kaufmann series in data management systems. Burlington, Mass.: Elsevier. <https://ebookcentral.proquest.com/lib/jyvaskyla-ebooks/detail.action?docID=729031>.

Isohookana, Heli. 2007. *Yrityksen markkinointiviestintä*. Toimittanut Heli Isohookana. Empiria. Lisäpainokset: 2. p. 2011. Helsinki: WSOYpro.

Kantardzic, Mehmed. 2011. *Data mining : concepts, models, methods, and algorithms*. 2nd ed. 534. Hoboken, New Jersey: John Wiley. <http://ieeexplore.ieee.org/xpl/bkabstractplus.jsp?bkn=6105606>.

Karjaluoto, Heikki. 2010. *Digitaalinen markkinointiviestintä : esimerkkejä parhaista käytännöistä yritys- ja kuluttajamarkkinointiin*. Jyväskylä: Docendo. <https://www.ellibslibrary.com/jyu/978-951-0-36091-0>.

Kotler, Phillip, Mairead Brady, Malcolm Goodman, Torben Hansen ja Kevin Keller. 2010. *Marketing Management*. Welwyn Garden City, UNITED KINGDOM: Pearson Education UK. <http://ebookcentral.proquest.com/lib/jyvaskyla-ebooks/detail.action?docID=5136334>.

Laveri, Jouni. 2006. "Käyttäjien profilointi verkkokaupoissa : viitekehysmalli profilointimenetelmien vertailuun". Tutkielma. <http://urn.fi/URN:NBN:fi:jyu-2006573>.

Leventhal, Barry. 2010. "An introduction to data mining and other techniques for advanced analytics". *Journal of Direct, Data and Digital Marketing Practice*, <https://doi.org/10.1057/dddmp.2010.35>. <https://doi.org/10.1057/dddmp.2010.35>.

- Linoff, Gordon S., ja Michael J. A. Berry. 2011. *Data Mining Techniques : For Marketing, Sales, and Customer Relationship Management*. John Wiley Sons, Incorporated. <http://ebookcentral.proquest.com/lib/jyvaskyla-ebooks/detail.action?docID=706770>.
- Merisavo, Marko, toimittanut. 2006. *Digitaalinen markkinointi*. Economica. Muut tekijät: Jari Vesänen, Mika Raulas ja Ville Virtanen. Helsinki: Talentum.
- Paytrail. 2021. *Verkkokauppa Suomessa 2020*. <https://www.paytrail.com/raportti/verkkokauppa-suomessa-2020>.
- Posti. 2020. *Suuri verkkokauppatutkimus 2020*. https://minun.posti.fi/hubfs/Tutkimukset/Suuri-verkkokauppatutkimus-2020_Posti.pdf#_ga=2.66485612.313805965.1617786317-780347764.1617786317.
- Päivittäistavarakauppa ry, PTY. 2021. *Päivittäistavarakaupan myynti ja markkinaosuudet 2020*. <https://www.pty.fi/ajankohtaista/tiedotteet/uutinen/article/paivittaeistavarakaupan-myynti-ja-markkinaosuudet-2020/>.
- Saura, Jose Ramon. 2020. “Using Data Sciences in Digital Marketing: Framework, methods, and performance metrics”. *Journal of Innovation Knowledge*, ISSN: 2444-569X. <https://doi.org/https://doi.org/10.1016/j.jik.2020.08.001>. <https://www.sciencedirect.com/science/article/pii/S2444569X20300329>.
- Tekin, Mahmut, Mehmet Etlioglu, Özdal Koyuncuoğlu ja Ertuğrul Tekin. 2019. “Data Mining in Digital Marketing”, 44–61. Tammikuu. ISBN: 978-3-319-92266-9. https://doi.org/10.1007/978-3-319-92267-6_4.
- Tuffery, Stéphane. 2011. *Data mining and statistics for decision making*. Toimittanut Stéphane Tuffery. Wiley series in computational statistics. Chichester, West Sussex ; Hoboken, NJ.: Wiley. <https://ebookcentral.proquest.com/lib/jyvaskyla-ebooks/detail.action?docID=792450>.
- Witten, I. H., Eibe Frank ja Mark A. Hall. 2011. *Data mining : practical machine learning tools and techniques*. 3rd ed. 629. Morgan Kaufmann series in data management systems. Burlington, MA: Morgan Kaufmann. <http://search.ebscohost.com/login.aspx?direct=true&scope=site&db=nlebk&AN=351343>.