

Adfore - Tutkimusteema: Personointi ja ennakoiva analytiikka osana kontekstietietoista kohdentamista, case Kontekstimoottori.

Selvitys apurahan käytöstä, loppuraportti

Tiivistelmä

Projektissa toteutettiin kolme eri casea, jotka painottuivat omiin osa-alueisiinsa.

1. case hyödynsi tekoälymenetelmiä, jotka rakentuivat käyttövalmiiden pilvipalveluiden päälle (Amazon Web Services). Menetelmillä kehitettiin sanomalehdille digitoitujen printti-ilmoitusten suosittelukone digiympäristöön. Tulokset olivat lupaavia vakioitujen perusratkaisujen toteuttamiseksi.

2. case keskittyi maakuntalehden asiakasymmärryksen lisäämiseen sekä asiakaspoistuman ennakointiin. Näitä varten kehitettiin kumpaankin tapaukseen räätälöidyt tekoälymallit. Mallien avulla pystyttiin tekemään uusia havaintoja asiakkaista mm. kohdennusta varten, ja ennakoimaan asiakkaiden käyttäytymistä tulevaisuudessa.

3. case keskittyi busseissa tapahtuvaan älykkääseen mainontaan. Paikka- ja muun kontekstiedon hyödyntäminen avasi uusia mainonnan mahdollisuuksia. Myös mobiiliin kytkeminen osaksi tarjoamaa laajentaa ja parantaa asiakaskokemusta. Personoinnin rooli oli tässä casessa pienempi. Painopiste oli sääntöpohjaisessa monipuolisessa mainosten kohdentamisessa eri kanavissa sekä mainonnan laajuuden ennakoinnissa.

Kaiken kaikkiaan projektissa saatiin lupaavia tuloksia tekoälyyn ja sääntökoneisiin perustuvien ratkaisujen hyödyntämisestä mediayhtiöiden eri palveluissa. Menetelmiä voidaan hyödyntää niin toimituksellisen aineiston kuin myös ilmoitusten osalta. Projekti loi vahvaa pohjaa tuottaa näihin ratkaisuihin perustuvia uusia palveluja mediayhtiöille. Kehitettyjen menetelmien lopullinen tuotteistaminen tapahtuu tämän tutkimusprojektin jälkeen.

Projektin aikana syntyi lisäksi myös ideoita toimituksellisen aineiston suosittelukoneen kehittämiseksi. Projektissa kokeiltiin artikkeleiden automaattista luokittelua text mining -algoritmien avulla, mutta luokittelua ei lopulta käytetty osana segmentointia. Luokittelutulokset antoivat kuitenkin osviittaa, että menetelmää (BlazingText¹) voisi käyttää artikkelikohtaisten suositusten tekemiseen.

Projektin tulokset

Case 1 – Personointi ja ennakoiva analytiikka mediayhtiön digipalvelussa – Tässä.fi (ja kaupunkilehdet)

Tässä casessa keskityttiin personoinnin kehittämiseen. Kansainvälisillä ja valtakunnallisilla medioilla on resursseja kehittää omia räätälöityjä menetelmiä personointiin ja ennakointiin, mutta maakuntatasolla tämä jo vaikeutuu, ja on käytännössä nykyisellään mahdotonta tätä pienemmillä mediayhtiöillä. Tästä syystä casen tavoitteeksi asetettiin löytää/kehittää mahdollisimman kustannustehokas menetelmä, joka voisi palvella myös pienempiä sanomalehtiä.

Personoinnin tavoitteeksi asetettiin käyttäjälle räätälöidyt sisältösuositukset. Sisältönä käytettiin Tässä.fi-palvelun ilmoitussisältöä sekä erityisesti palvelussa mukana olevien kaupunkilehtien digitaaliseen muotoon muutettuja printti-ilmoituksia.

Amazon julkaisi kesäkuussa omassa pilvipalveluympäristössään uuden Personalize-palvelun, joka tarjoaa Amazonin itse verkkokaupassaan käyttämiä suositteualgoritmeja kenen tahansa käyttöön. Palvelu on erittäin pitkälle tuotteistettu eikä sen käyttöönotto sekä käyttö vaadi erityistä tekoälyalgoritmien osaamista. Liiketoimintamalliltaan palvelu näytti sopivan hyvin tässä casessa haettuun kustannustehokkaaseen ratkaisuun. Omien tekoälymallien tekeminen ja ylläpitäminen yksittäisille kaupunkilehdille olisi liian raskasta kustannusten puolesta, joten keskitetty ja palveluna ostettava ratkaisu vaikutti hyvältä.

Integraatiot Tässä.fi-palvelun ja Personalize-palvelun² välillä tehtiin kesän ja alkusyksyn aikana. Personalize-palvelu vaatii hyvin määrämuotoista dataa, ottaen vastaan käytännössä kolmea eri tietotyyppiä:

- Interaktiot. Tämä on tärkein datasetti, josta ilmenee mitä sisältöä kukin käyttäjä on kuluttanut.
- Tuotteet. Metatietoa tuotteista eli tässä casessa ilmoituksista. Tärkein metatieto oli ilmoitusten luokitus (esim. ruokakauppa, lounasravintola, parturikampaamo).
- Käyttäjät. Metatietoa käyttäjistä.

Palvelu saatiin toimimaan syyskuun 2019 aikana. Personoinnin osalta toteutettiin kaksi eri käyttötapausta:

1. Ilmoitusten suositteelu perustuen mitä muut samoja ilmoituksia katselleet käyttäjät ovat katsoneet (*collaborative filtering*).
2. Sanomalehden kaikkien printti-ilmoitusten järjestäminen käyttäjäkohtaiseen suositusjärjestykseen. Tässä tapauksessa käyttäjälle voidaan personoida oma ilmoitusnäkyminen hänen käyttöhistoriansa perusteella niin, että sanomalehden ilmoitukset näytetään hänelle räätälöidystä järjestyksessä (*content-based filtering*).

Tuloksien mittaamiseen ei pystytty muodostamaan kvantitatiivisia mittareita, joten tuloksia on arvioitu projektiryhmän sisällä oma-arviona. Suositukset näyttävät toimivan loogisesti ja sitä paremmin, mitä enemmän klikkidataa algoritmilla on käytössä.

Varjopuolena valitussa lähestymistavassa on menetelmien jatkokehitys. Tällainen palveluna tarjottava tekoälymalli on kuin musta laatikko, jonka sisälle ei pysty tekemään kovinkaan paljon muutoksia. Menetelmät ovat valmiita ja ennalta määrättyjä, ainoastaan tiettyjä parametreja voi muuttaa. Osoittautui kuitenkin ilmeiseksi, että jo parametrien muuttamiseen ja seurausten analysointiin tarvittaisiin tekoälymenetelmien parempi tunteminen (data scientist).

Yhteenvetona voidaan todeta, että Amazon Personalize ja vastaavat muut mahdolliset black box -tekoälymallit voivat toimia silloin, kun tarvitaan ratkaisua hyvin yleiseen ongelmaan ja käytettävät tekoälymallit ovat riittävän hyviä laatutasoltaan. Tällainen palvelu on mahdollista saada käyttöön jopa vain muutaman sadan euron kuukausimaksulla, joten pienille toimijoille se voi olla ainoa mahdollinen ratkaisu. Adforen Tässä.fi-palvelussa malli näyttää toimivan riittävän hyvin ja tarkoitus on laajentaa sen käyttöä ja markkinoida ratkaisua myös mukana oleville kaupunkilehdille. Palvelua ei testattu uutisaineistolla, mutta se voisi toimia siinäkin. Tässä tapauksessa haasteeksi saattaa muodostua uutisten lyhyt elinkaari, sillä mallien opetus ja suositusten saaminen uusista uutisista vie jonkin verran aikaa.

Case 2 – Personointi ja ennakoiva analytiikka mediayhtiön digipalvelussa (maakuntalehti)

Casen vaatimusmäärittelyssä päädyttiin siihen, että kuluttajakäyttäytymisen ymmärtäminen ja ennakointi ovat keskeisiä kehittämisasioita. Näistä lähtökohdista ilmeni heti, että edellisessä casessa käytettyjä menetelmiä ja palveluja ei voinut soveltaa tässä tapauksessa. Vaatimusmäärittelyssä päädyttiin kahteen osa-alueeseen:

1. Asiakaskäyttäytymisen ymmärtäminen RF-analyysin* sekä lukutottumuksiin perustuvan tekoälypohjaisen segmentoinnin avulla.
2. Asiakaspoistuman ennakointi

RF-analyysi antoi selkeän kuvan, miten asiakkaat jakautuvat eri segmentteihin perustuen käynti- ja lukuaktiivisuuteen. Pelkästään tällä melko yksinkertaisella mallilla on mahdollisuus lisätä yrityksen omaa asiakasymmärrystä. Silti haasteeksi jää se, että minkälaisia esim. hiipuvat asiakkaat ovat ja minkälaisilla viesteillä tai sisällöillä heitä kannattaisi aktivoida. Tähän liittyen projektissa tehtiin tekoälyyn perustuva segmentointi lukukäyttäytymisen perusteella.

Kuluttajakäyttäytymisen ymmärtämisessä keskeinen asia on tutkia käyttäytymismalleja ja segmentoida käyttäjiä eri käyttäjäryhmiin. Saatuja tuloksia voidaan suoraan hyödyntää monella tavalla, esim. personoimalla markkinointia tai suunnitella uutistuotantoa kohderyhmänäkökulmasta.

Segmentoinnissa päädyttiin tekemään ”puhdas” artikkelien lukemiseen perustuva luokittelu. Näin haluttiin toisaalta ymmärtää, mistä aiheista eri käyttäjäryhmät ovat kiinnostuneet, mutta myös tuomaan lisätietoa lehden toimitukselle siitä, minkälaisia juttuja näille eri ryhmille kannattaisi kirjoittaa.

Segmentointi tehtiin hierarkisena klusterointina Wardin menetelmällä. Wardin menetelmä minimoi klustereiden hajonnan ja muistuttaa paljon käytetyn k-means menetelmän kohdefunktiota. Päätason segmenttejä päätettiin muodostaa kolme ja alemman tason segmenttejä yhdeksän. Etenkin näillä yhdeksällä segmentillä ajateltiin saatavan riittävän selkeät eri profiiliset segmentit, joiden avulla toimituksen ymmärrys asiakkaista kasvaa ja toisaalta personointi ja markkinoinnin kohdentaminen onnistuu hyvinkin tarkasti. Käytimme KNN-luokitinta yleistämään hierarkisen klusteroinnin tulokset koko populaatioon. Eli alkuperäinen klusterointi tehtiin vain osalla asiakkaista, mutta klusterointitulokset haluttiin saada kaikille.

Poistuman ennakoinnissa käytettiin Random Forest -menetelmää³ (päättöpuumenetelmiin pohjautuva epälineaarinen yhdistelmäluokitin). Menetelmän avulla saatiin luotua kohtuullisen toimiva malli poistuman ennakoimiseen. Käytössä oleva data asetti rajoitteita, esim. poistuneiden asiakkaiden määrä oli melko pieni.

Malli pystyy opetusaineistolla ennustamaan palvelusta poistuvat asiakkaat melko tarkasti. Mallin mahdollisena heikkoutena on ”väärät positiiviset”.

* RF-analyysi. Recency-frequency-analyysi luokittelee käyttäjät eri luokkiin viimeisen käyntikerran (recency) perusteella sekä toisaalta aktiivisuuden (frequency) perusteella. Useissa tutkimuksissa on todettu näiden tekijöiden vaikuttavan merkittävästi kuluttajien ostokäyttäytymiseen. Malliin liitetään usein vielä yksi dimensio (Monetary, RFM), joka kuvaa asiakkaan ostosten arvoa. Tilauspohjaisessa liiketoimintamallissa tämä dimensio ei ole kovin kuvaava ja on jätetty tarkastelusta pois.

Esim. Fader, P. S., Hardie, B. G., & Lee, K. L. 2005. RFM and CLV: Using iso-value curves for customer base analysis. *Journal of Marketing Research*, 42(4), 415-430. ja Khajvand, M., Zolfaghar K., Ashoori, S., Alizadeh, S., 2011. Estimating Customer Lifetime Value Based On RFM Analysis of Customer Purchase Behavior: Case Study, *Procedia Computer Science*, Vol. 3, pp.57-63.

Case 3 – Personointi ja ennakoiva analytiikka hyperlokaalin ulkomainonnan ja mobiilin yhdistelmänä

Erikoisuutena tässä casessa oli ulkomainonnan (DOOH, Digital Out Of Home) näyttöjen eli mainospinnan jatkuva liikkeellä olo (näytöt linja-autoissa). Tämä sijainnin vaihtelu mahdollistaa myös uudenlaiset mainonnan muodot. Projektissa kehitettiin järjestelmä, jossa Adforen Kontekstimoottorille päivitetään jatkuvasti reaaliaikaisena bussien sijaintitiedot. Kyseessä on IoT eli Internet of Things -ratkaisu, jossa gps-paikantimet busseissa välittävät paikkatiedon Adforen palvelimille. Kontekstimoottoriin luotiin valmiudet toteuttaa mainontaa erilaisten sääntöjen pohjalta. Lisäksi testattiin teknisesti, miten samaa mainontaa voidaan ohjata myös mobiililaitteeseen. Jyväskylä oli pilottikohteena ja Jyväskylän busseihin asennettiin myös beacon-lähettimet. Näiden lähettimien avulla puhelimesta oleva sovellus kykenee tunnistamaan, milloin käyttäjä on bussissa. Tietosuoja (GDPR) on aina käyttäjäsovelluksen vastuulla, sillä beaconit ovat vain ”tyhmiä” bluetooth-lähettimeitä, jotka lähettävät signaalia niitä kuunteleville puhelimille/sovelluksille, eivätkä itsessään kerää käyttötietoa.

Kontekstimoottori kytkettiin näyttämään Adforen Tässä.fi-sovelluksessa samaa mainontaa kuin bussien näytöillä esitettiin. Tämä toteutettiin kontekstikorttien avulla. Lisäksi testattiin, että puhelimesta tapahtuvaa mainontaa voidaan edelleen personoida käyttäjän profiilin mukaisesti. Esimerkiksi Jyväskylässä paikalliselle jyväskyläläiselle henkilölle voidaan näyttää erilaista mainontaa puhelimesta kuin pistäytymässä olevalle vieraspaikkakuntalaiselle. Puhelimesta näkyvä mainonta auttaa myös call-to-actionin toteutumista eli että kuluttaja voi esim. siirtyä mainostajan mainostamalle verkkosivulle ja katsoa lisätietoja, vaikkapa myynnissä olevasta asunnosta.

Kesällä toteutettiin pilotointi, jossa testattiin Kontekstimoottorin toimivuutta paikkatietoisessa ja sääntöpohjaisessa mainonnassa. Pilottiasiakkaita oli kaksi, OP-Koti ja Keskimaa. OP-kodin tapauksessa busseissa ja puhelimesta näytettiin aina lähintä myynnissä olevaa asuntoa. Sisältö siis vaihtui dynaamisesti bussien liikkua. Pilotoinnin tulokset olivat onnistuneita, tekniikka toimi hyvin. Alla kuva mainonnasta:



Osuuskauppa Keskimaan kanssa toteutettiin säännöiltään monimutkaisempi pilotti. Kesällä pilotoitiin ratkaisua, jossa mainostettiin bussia lähinnä olevaa osuuskaupan ketjua. Mainonta vaihteli siis ravintoloiden, ABC-asemien sekä ruokakauppojen välillä. Lisäksi mainonnalle asetettiin myös muita sääntöjä, joissa viikonpäivä, vuorokaudenaika sekä jopa sää vaikutti mitä näytetään. Esimerkiksi Jyväskylän keskustassa sijaitsevan ravintolan mainosta näytettiin tiettyinä viikonpäivinä ja tiettyinä kellonaikoina, mutta vain jos aurinko paistoi eli oli terassikeli.

Kontekstimoottori saatiin toimimaan hyvin näiden jo melko vaativienkin mainonnan kohdentamisen sääntöjen osalta.

¹ Saurabh Gupta and Vineet Khare. 2017. BlazingText: Scaling and Accelerating Word2Vec using Multiple GPUs. *Proceedings of the Machine Learning on HPC Environments (MLHPC'17)*. ACM, New York, NY, USA. DOI: <https://doi.org/10.1145/3146347.3146354>

² Real-time personalization and recommendation | Amazon Personalize | AWS
<<https://aws.amazon.com/personalize/>>. 17.12.2019.

³ Ho, Tin Kam. 1995. Random Decision Forests. *Proceedings of the 3rd International Conference on Document Analysis and Recognition*, Montreal, QC. DOI: <https://doi.org/10.1109/ICDAR.1995.598994>