

# 'Big Data Analytics' i Telenor Group

**I Telenor Group Research har man satt opp et team av 'data scientists' hvor det ligger i mandatet å kunne forske, utvikle og teste nye måter å benytte kundedata på. Dette er et tverrfaglig team satt sammen av personer med bakgrunn i numerikk-tunge fagretninger som matematikk, fysikk og data science, men også mykere fag som psykologi og sosiologi. Fellesnevneren er stor interesse og høy kompetanse innen numerisk analyse og behandling av store datamengder. I tillegg til den interne kompetansen, samarbeider Telenor Research med 'big data' forskningsmiljøer på Harvard, MIT, Cambridge og Northeastern University. Gruppen fungerer som brobyggere mellom business og akademia og har fokus på å bringe forskningsresultater inn i forretningsvirksomheten. Telenor Research har over flere år bygget opp en unik forskningsdatabase med anonymiserte kundedata fra flere land.**

TEKST: JOHANNES BJELLAND, johannes.bjelland@telenor.com OG PÅL SUNDSØY, pal-roe.sundsoy@telenor.com



Pål Sundsøy.

Johannes Bjelland.

## Bakgrunn

'Big Data' er et begrep som har klatret fort opp på Gartners hype-kurve for emerging technologies de senere år. Big Data har også fått mye omtale i media, spesielt i forhold til internetselskaper som Facebook, Google og Amazon. Vi i Telenor Research Data Analytics gruppe mener det har vært mye «buzz» og få konkrete eksempler på hvordan disse dataene faktisk kan brukes til noe nyttig. Forrige uke ble også Gartner's kurve for 2014 lansert – her er 'big data' godt på vei ned i 'trough of disillusionment', og er generelt et godt stykke unna å generere inntekter. Det mange ofte glemmer når man tenker på 'Big data', er at Telekom-bransjen også sitter på store mengder data om sine kunder. Daglig pumpes det inn titalls milliarder rader trafikkdata inn i databasene til Telenor. Disse dataene blir i dag stort sett brukt til å generere fakturaer. Vi vet at det finnes et stort ubenyttet potensial i disse dataene, og vi mener at Telekombransjen bør bli flinkere til å benytte denne kilden til innsikt.

## Hva telekomdata kan fortelle oss om menneskelig oppførsel

Mobiltelefonen er blitt en del av oss. For mange er den innen rekkevidde 24 timer i døgnet, og i mange land bruker så godt som 100% av befolkningen en mobiltelefon. Dette gjør telekom-data til en revolusjonerende ny kilde til storskala studier av menneskelig oppførsel.

Et av de viktigste forskningsområdene der telekom data har bidratt, er utvikling av kunnskap om hvordan våre sosiale nettverk er bygd opp og hvordan det påvirker vår oppførsel. Fra vår database med anonymiserte ringe-logger kan vi konstruere et kommunikasjonsnettverk. Med noe filtrering er et slikt nettverk er en god tilnærming til det virkelige sosiale nettverket – vi ringer våre venner. Vi har blant annet sett på hvordan produkter spres viralt over sosiale nettverk og har dermed kunnet kvantifisere jungeltelegrafens betydning i markedsføring. Et enkelt resultat viser at dersom din venn har en iPhone, så ti-dobles sannsynligheten for at du vil skaffe deg en også. Et fascinerende bilde av denne effekten vises i Figur 1. Her ser vi hvordan iPhone spredte seg som ild i tørt gress og til slutt danner et ekstremt tett nettverk. Vi jobber med marketing-team for å kunne benytte denne innsikten i fremtidige kampanjer.

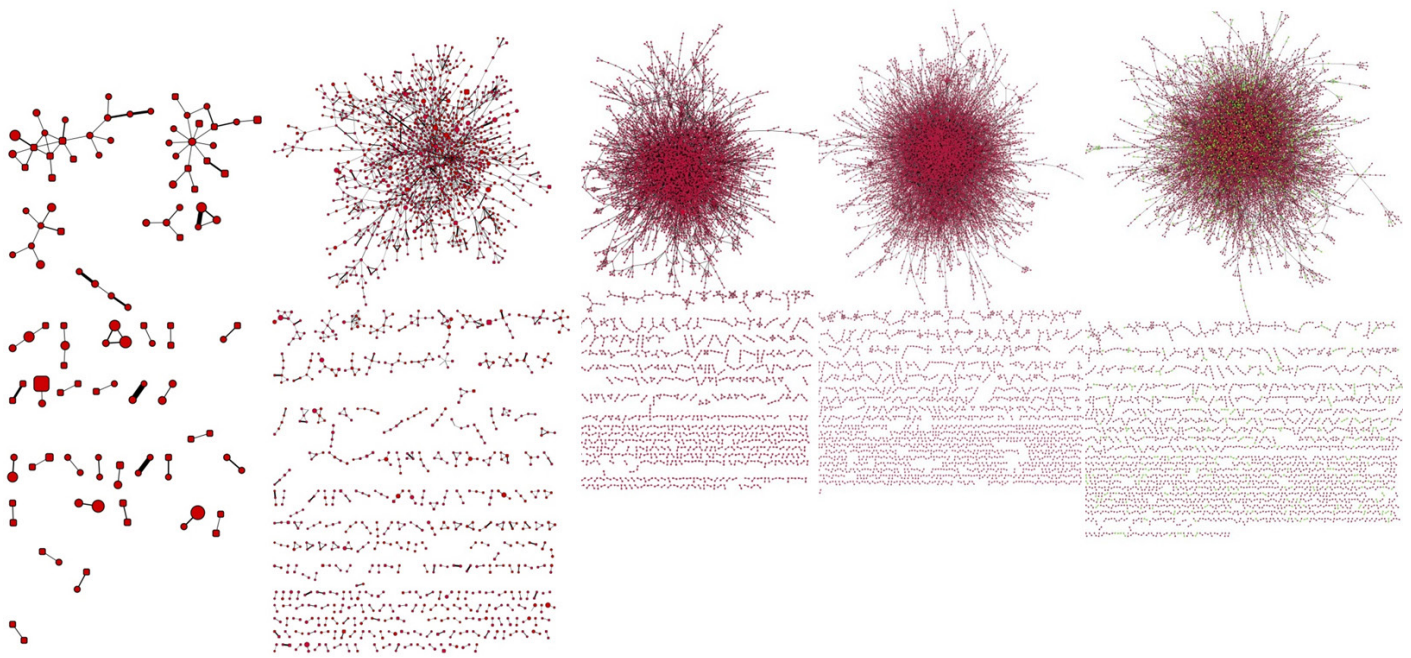
Et annet aspekt ved Telekom data er at vi kan bruke informasjon om belastning på ulike basestasjoner for å måle hvordan store

folkemengder forflytter seg geografisk. Slike analyser har vist seg å være nyttige i byplanlegging, trafikkanalyser og effektivisering av infrastruktur. Som eksempel ser vi i Figur 2 hvordan basestasjonsbelastningen fordeler seg under 17. mai i Oslo.

## Effektiv markedsføring med Big Data Analytics gir 13 ganger bedre treffrate

For å tallfeste verdien av en data-drevet fremgangsmåte i forhold til klassisk metode, satte vi opp et eksperiment i et av Telenors Asiatiske markeder. Casen vi bruker er markedsføring av internett via mobil. I dette markedet er internettpenetrasjonen lav og mobiltelefonen er ofte eneste mulighet for mange å komme på nett. En stor del av kundene har også en telefon med mulighet for databruk, men vi ser at dette i mange tilfeller ikke blir brukt. Dermed er potensialet for å konvertere kunder til databrukere stort. 'Internet for All' er også et motto og et stort satsningsområde for Telenor Group i Asia. For også å sikre god akademisk relevans og kvalitet, samarbeider vi med Big Data forskningsmiljøet ved Northeastern University og MIT's MediaLab. Dette er noen av de fremste internasjonale ekspertene på området.

I de fleste av våre asiatiske markeder bruker telekomselskaper mye ressurser på direkte markedsføring via SMS. Det er en billig og enkel måte å nå sine kunder på. Problemet er at det har vært for billig og for enkelt og vi ser at det er en tendens til overbruk. Dette er

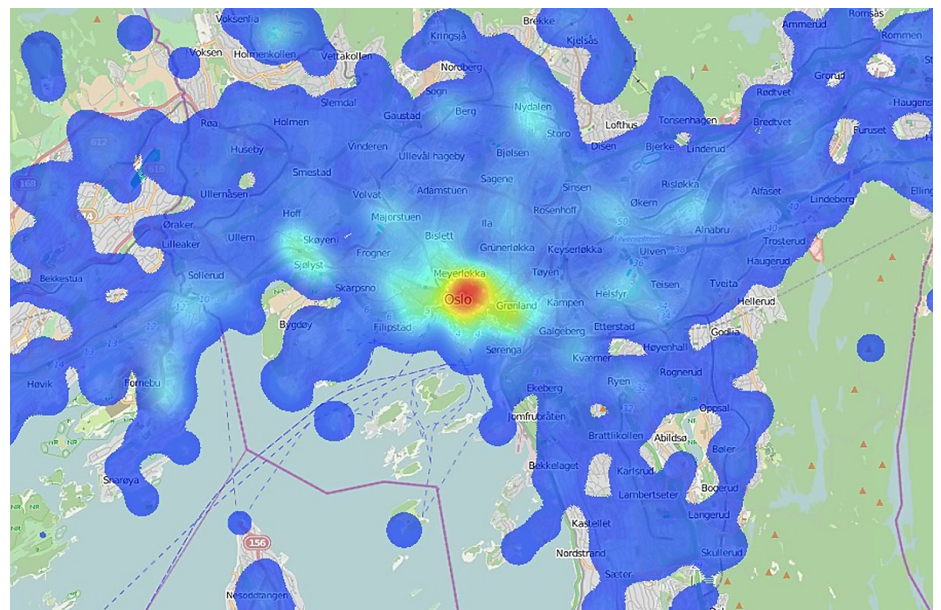


Figur 1. Evolusjonen av det sosiale nettverket mellom iPhone-brukere over tid. Hver prikk representerer en kunde, og fargen representerer hvilken iPhone modell som er i bruk.

ikke bare irriterende for kunden, vi er i ferd med å ødelegge SMS som salgskanal: Kundene ser an avsender og gidder ikke engang å åpne meldinger fra oss. Dette bekreftes av studier som viser at mindre enn 30% av kampanje-meldingene blir lagt merke til og lest av kunder.

Som regel sendes slike tilbud ut basert på enkle analyser, som typisk involverer antagelser og markedsførernes magesfølelse om hvem som kan være interessert i produktet. En nylig studie utført av IBM viser at 80% av de spurte markedsførerne har gjort kampanjebeslutninger basert på slike antagelser. Typisk vil man for eksempel anta at dette produktet er interessant for et yngre alderssegment eller kunder med lavt SMS forbruk. Dette er hva vi kaller den klassiske metoden som vi vil sammenligne mot. I stedet for å la mennesker velge ut segmentet, vil vi la Machine Learning (ML) algoritmer gjøre jobben med å velge ut hvilke kunder som skal motta dette tilbudet.

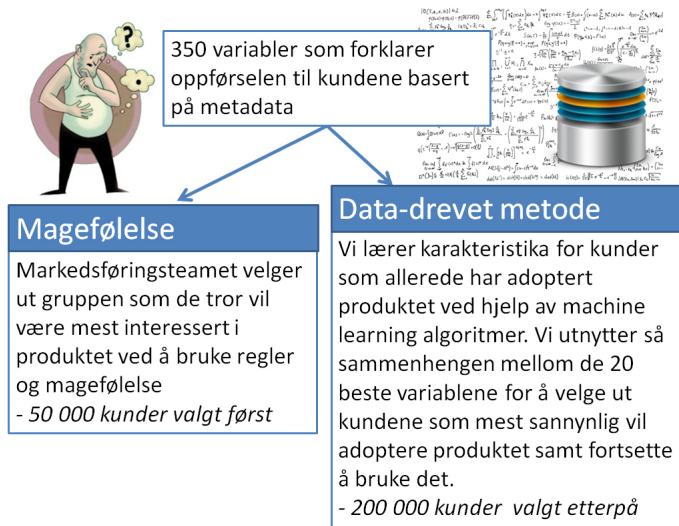
På hver kunde har vi registrert over 350 ulike attributter. Disse variablene inneholder for eksempel abonnemehistorikk, forbruksmønstre og egenskaper ved håndsettet. I det aktuelle markedet er det ca. 60 millioner kunder. Med over 350 variabler pr kunde ender man opp med et stort og komplekst datasett som skal analyseres for å finne de skjulte mønstrene som identifiserer den optimale målgruppen. Uten spesialsoftware og machine learning algoritmer ville dette



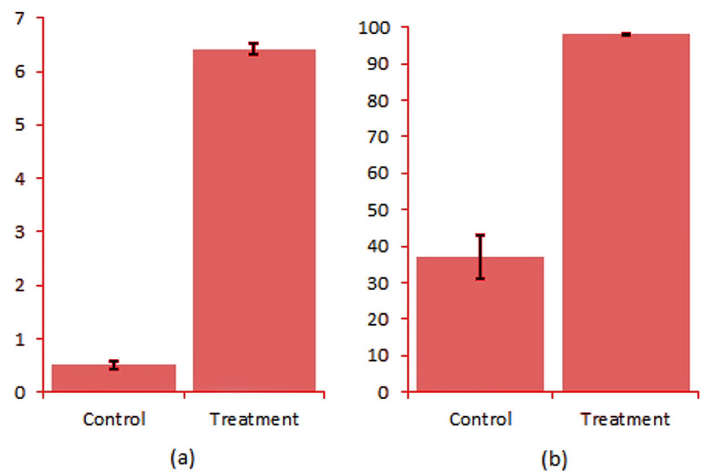
Figur 2. Bildet illustrerer «trøkket» på basestasjonene i Oslo 17.mai i år. Man kan observere forløpet av bametoget og aktiviteten i parkene. Animasjon tilgjengelig i Dagens Næringsliv: <http://goo.gl/23zknc>.

*Et av de viktigste forskningsområdene der telekom data har bidratt, er utvikling av kunnskap om hvordan våre sosiale nettverk er bygd opp og hvordan det påvirker vår oppførelse.*





Figur 3. Sammenligning av standard markedsføringsmetoder VS data-drevne prediksjonsmetode.



Figur 4. Data-drevet metode gir 13 ganger bedre treffrate. (a) Konverteringsrate for kontroll (best-practice) og treatment (data-driven approach) (b) Prosent kunder som tegnet ny dataplan etter kampanjen («renewal rate»). For flere detaljer se vitenskaplig publikasjon (medforfattet med MIT): <http://goo.gl/OLiV2>

ikke være mulig. Så store datamengder krever at vi tar i bruk big data metoder.

Som kontroll lar vi først teamet av erfarne markedsførere velge sin målgruppe, deretter lar vi algoritmene velge. Totalt sender vi ut 250 000 sms med et godt tilbud på mobilt internetturfing og registrerer svar.

Resultatet er overbevisende: Marketing teamet oppnår en konversjonsrate på 0.5% mens algoritmene konverterer 6.4%. Algoritmer og big data metodene er altså i dette tilfellet nesten 13 ganger så effektive til å velge ut relevante målgrupper.

Algoritmen ble optimalisert til å unngå såkalte «free-riders» – kunder som kun er interessert i gratis og tilbudskampanjetilbud. For å studere denne effekten ser vi på andelen av kunder som takket ja til det første tilbudet og som deretter fortsetter over på fullpris-alternativet når kampanjeperioden er over. Algoritmene vinner igjen: 98% av kundene valgt av ML-metoden fortsetter over på fullpris-alternativet, mot 37% i gruppen valgt av markedsførerne (se figur 4).

Dette eksperimentet er et eksempel på hvordan man kan bruke kundedata til å finne den optimale mottaksgruppen for en kampanje. Som regel har man ti- til hundretalls kampanjer kjørende samtidig. For virkelig å høste gevinsten av big data metoden og redusere unødig «spamming» av egne kunder til et minimum, bør man bruke denne metoden i stor skala. Vi kan i teorien lage en slik modell for alle typer av produkter og kampanjer. Da har vi også

mulighet til å optimalisere rett produkt til rett kunde – Det vil si dersom en kunde har mulighet til å velge både produkt X og produkt Y, så kan vi la algoritmen avgjøre ved å velge produktet som er mest relevant for akkurat denne kunden.

**Big data er ikke bare business: Analyser for en bedre verden**

I tillegg til å forbedre kommunikasjon og anbefalinger til kunden, så kan telekommunikasjonsdata brukes av forskere som utvikler modeller for smittespredning. Ved å se på hvordan ringevolum fordeles på basestasjoner kan man estimere hvordan menneskemasser forflytter seg innad i et land. Siden smitte følger menneskers bevegelse er dette et veldig viktig parameter i modeller for spredning av sykdommer. Telenor Research jobber sammen med forskere fra Harvard University med å se på hvordan våre data kan bidra til å forbedre sykdomsspredningsmodeller i utviklingsland.

Dersom man kan bli bedre på å forutsi et smitteforløp, kan man enklere planlegge vaksinasjonsprogram og distribusjon av medisiner. Det samme gjelder også ved store naturkatastrofer som tsunamier og jordskjelv: Hvor skal man sende nødhjelp og hvor skal man bygge leirer for husløse flyktninger? Ringevolumer aggregert på geografisk nivå kan hjelpe med planlegging/distribusjon av nødhjelp og kun en liten forbedring vil kunne redde menneskeliv. Visualiseringsteknologien som viser bevegelse av folkemengder 17. mai kan like så godt være nyttig ved humanitære katastrofer.



Figur 5. Telenor opererer i flere markeder der Big Data analyser kan brukes for å skape en bedre verden. Foto fra Myanmar: Pål Sundsøy

*Resultatet er overbevisende: Marketing teamet oppnår en konversjonsrate på 0.5% mens algoritmene konverterer 6.4%.*