

(c) Prof.dr.sc. Goran Klepac

goran(..at..)goranklepac.com

Kad anketari utihnu

U hladno zimsko predvečerje ...

„Konačno, petak poslijepodne. Neopterećen brigama sjedam pokraj kamina, uronjen u toplinu sobe dok se izvana čuje fijuk vjetra. Bilo bi u stvari lijepo raditi od kuće, umjesto svaki dan odlaziti na posao, raditi od 10:00 sati umjesto od 8:00. Buditi se sa okusom prave kave u ustima a ne one iz automata. Uostalom kada bolje razmislim, ova kava 'Zrno' koju upravo kuham je baš po mom ukusu, a kada bi mi je još pri ruci bio i novi CD Diane Krall, doživljaj bi bio potpun... “

Pretpostavimo kako je ovo isječak iz bloga. Gledajući iz globalne, svjetske perspektive, dnevno se u cyber prostor putem blogova slije more tekstova sličnog karaktera. Neki od tekstova kreirani su u svrhu terapijskog učinka, jer primjerice njegov kreator/kreatorica upravo proživljava dramatične emotivne trenutke u svom životu. Neki od tekstova vezani su uz stručnu problematiku jer je netko upravo u domeni svoje struke otkrio značajnu činjenicu te to želi podijeliti s nekim. Bez obzira na karakter tekstova, svaki od njih često je puta prožet određenim osobnim elementima, odnosno elementima koji nisu direktno vezani uz razmatranu problematiku. Ljudi u tim tekstovima otkrivaju sebe i što je najvažnije, zaštićeni pseudonimima ne kalkuliraju informacijama koje ostavljaju u cyber prostoru već te informacije nose značajnu dozu vjerodostojnosti.

Pretpostavimo kako je poduzeće 'Pržionica kave', upravo lansiralo novu vrstu kave 'Zrno', te je zainteresirano istražiti kako je tržište prihvatilo njihov novi proizvod koji cilja specifični tržišni segment.

Kako izraziti impresiju na skali od 1 do 10 ?

Tradicionalni pristup sveo bi se na klasično istraživanje tržišta putem telefonskih anketa. Tko od vas nije doživio poziv poslije radnog vremena i monoton glas s druge strane telefonske linije koji vas pita što mislite o proizvodu X i dali ste u posljednje vrijeme primijetili reklamni spot na televiziji koji govori o tom proizvodu ? Umorni od napornog radnog dana vi isto tako mehanički odgovarate na šablonizirana pitanja, uz uvjet da prije toga niste odbili sudjelovati u anketi. Ako pak slučajno imate dodatan komentar na primjerice obavljen servis vašemeg vozila, glas sa druge strane ljubazno vas moli ako svoju impresiju možete izraziti isključivo skalom od 1 do 10 bez dodatnih komentara. Komentar se možda svodio na dobru uslugu, ali možda na hladan prijem u autoservisu ili pak na nedovoljnu zainteresiranost majstora obraćanja pažnje na problem koji je vama bitan, a njemu se to činilo nevažnim. Sa druge strane, postavlja se pitanje da li ste upravo vi prava osoba koju treba nešto pitati o nekom proizvodu ili usluzi, i da li vaši odgovori doprinose kvaliteti provedenog istraživanja.

Kako reklamne kampanje utječu na podsvijest

Vratimo se sada blogu s početka priče. Iz isječka bloga naziremo životni stil pisca bloga, određen stav, težnje... Iako se čini nevjerovatno, ali iskustvo (USA) je pokazalo kako se na dovoljno velikom uzorku blogova određeni proizvodi (novolansirani) i tematike spominju podsvjesno i u blogovima ovisno o intenzitetima reklamnih kampanji. Ti elementi su utkani u određenom kontekstu unutar tekstova gdje bi se tome najmanje nadali. Tako ih možemo pronaći u negativnom ili pozitivnom kontekstu u stilu: "...trčao sam preko livade poput onog lika u reklami za" ... što upućuje na zamjećivanje reklame. Isto tako rečenica tipa : "... iritantna poput proizvoda X ..." upućuje na stav o nekom proizvodu.

Ako postoji statistički značajan broj slučajeva koji pokazuju iste ili slične karakteristike, tada se ta tvrdnja smatra relevantnom i odražava stav određene skupine o određenoj problematici, proizvodu ili usluzi.

Odjek krampova u virtualnim prostorima

To ukratko znači kako postoji kritična masa blogova koji se kroz tekstove možda negdje referiraju na određeni proizvod ili uslugu. Sa druge strane iz tih tekstova moguće je spoznati životne stilove pisaca tih blogova. Kategorizirati ih prema stavovima njihovih autora, sustavima vrijednosti i slično. Naravno nitko ne pomišlja kako je za uspješnu realizaciju tih aktivnosti potrebno pročitati tisuće, desetine tisuća, milione stranica tekstova. Tu nam pomaže informacijska tehnologija. Kako je ovdje riječ o nestrukturiranim podacima (tekstu), rješenje moramo tražiti u sferama tekst mininga. Tekst mining je pod-područje data mininga, koje se bavi otkrivanjem znanja iz tekstualnih datoteka.

Današnje metode text mininga su značajno odmakle od metodologije pukog prebrojavanja frekvencija pojavnosti ključnih riječi i njihovih poveznica, već su u određenoj mjeri sposobne analizirati i kontekst sadržaja.

Zlatni standard

Proces učenja iz teksta započinje definiranjem zlatnog standarda. To je korak u kojem je presudan ljudski faktor i gdje određeni broj ljudi analizira i čita kritičan broj reprezentativnih blogova, kako bi mogao mapirati ključne pojmove definirati apstraktne nadkategorije i spoznati osnovne karakteristike kreatora ciljne skupine blogova.

S tim u skladu definiraju se se nadkategorije koje karakterizira skup ključnih riječi i prepoznavanje konteksta sadržaja. Tako se primjerice tekstovi mogu kategorizirati prema kontekstu. Trenutno je jedan od najvećih problema sa kojima se suočavaju modelari ovakvih sustava prepoznavanje ironije iz konteksta. Analitičari se u svom poslu mogu služiti i "trikovima" analizirajući učestalost korištenja znakova poput ":", "(", i slično s obzirom na dob, spol i ostale karakteristike u procesu mapiranja, odnosno definiranja zlatnog standarda. To znači da analizom učestalosti korištenja ovih znakova unutar blogova možemo reverzibilno procijeniti dob, spol, odnosno ostale karakteristike. Stvari su mnogo prizemnije primjerice u nastojanju prepoznavanja spola kroz blog, prepoznavajući uzorke poput "moj aparat za brijanje" ili pak prepoznavajući zaokupljenost autora bloga karakterističnom tematikom.

Plemena u cyberprostoru

Prilikom analiza blogova koristi se i tehnika definicije "plemena". Metodologiju "plemena" karakterizira prepoznavanje skupine blogera istih karakteristika, fokusa interesa ili zaokupljenosti određenom problematikom, te se tako prepoznate skupine sustavno analiziraju u jedinici vremena. Prate se njihove aktivnosti, trenutne zaokupljenosti, motivacije želje i slično.

Zamislimo tako prepoznato "pleme" hobista plastičnog maketarstva, gdje njihovi pripadnici redovito putem blogova objavljuju svoje radove, iskazuju stavove, recenziraju nove proizvode, iznose svoje namjere o budućim projektima, kao i kritike na račun postojećih proizvoda. Ako je pak dio pripadnika tog "plemena" prisutan i na socijalnim mrežama, postoji i više nego solidna infrastruktura za naprednu analizu i marketinške aktivnosti.

„Plemenski“ pristup u cyberprostoru koristi se i prilikom kreiranja i osmišljavanja izbornih kampanja.

Zamislimo tek koliko bi informacije sakupljene na ovaj način bile od koristi našim političarima u jeku predizbornih kampanji? Ili bi ih možda previše izdeprimirale?

Praćenje stavova i trendova u ovoj skupini može rezultirati idejama za razvoj novih proizvoda, poboljšanje postojećih proizvoda, kao i promjenama postojećih strateških odrednica. Naravno, da bi se u potpunosti mogli osloniti na ovako ekstrahirane informacije, potrebno je ekstrahirati informacije iz dovoljnog uzorka blogova, kao bismo bili sigurni u pouzdanost tako provedene analize.

U USA postoje specijalizirane agencije, koje se bave ovakvom vrstom analiza/istraživanja tržišta. Njihov raspon usluga seže od prepoznavanja određenih dominantnih skupina "plemena" u cyber prostoru, do fokusiranja na skupine određenih karakteristika ili pak korisnika određenih proizvoda i usluga.

Takve agencije objavljuju i periodične publikacije u kojima analiziraju trendove određenih skupina, te daju prognoze za buduća razdoblja.

Neočekivani tržišni segment

Koristeći ovu metodologiju, jedna američka kompanija koja se bavi proizvodnjom i distribucijom energetske pića došla je do neočekivanog otkrića, koje je rezultiralo otvaranjem novog tržišnog segmenta korisnika njihovih proizvoda. Naime, nakon lansiranja nove vrste energetske pića namijenjenog polaznicima fitnessa, angažirali su agenciju koja se bavi analizom blogova. Cilj istraživanja bio je otkriti kako su polaznici fitnessa prihvatili novi proizvod koji bi im trebao osigurati veću količinu energije tijekom vježbi, i da li je ovaj tržišni segment percipirao novi proizvod kroz reklamne kampanje. Analizirajući ciljani tržišni segment, tijekom relativno kratkog vremenskog perioda pojavio se neočekivani, dijametralno suprotni uzorak korisnika spomenutog energetske pića. Taj su uzorak u značajnoj mjeri činili studenti skloni noćnim izlascima i prekomjernom uživanju alkohola. Empirijski su spoznali, konzumirajući napitak tijekom noći nastojeći što duže ostati budni, kako on ublažava mamurluk, nakon pretjeranog uživanja u alkoholu. To je rezultiralo nizom svjedočanstava kroz blogove i potaknulo ostale pripadnike ovog „plemena“ u iskušavanju preporučenog recepta. Kompanija, obogaćena novim spoznajama fokusirala se je i na ovaj tržišni segment sa izmijenjenom strategijom tržišnog komuniciranja prilagođenom upravo ovom „plemenu“, što je rezultiralo povećanjem prodaje.

Šumovi u kanalu

Ekspanzijom blogova, pojavom agencija koje analiziraju blogove i javno objavljuju rezultate takvih analiza, pojedinci, a više poduzeća došli su na ideju kreiranja sblogova odnosno spam blogova. Osnovna karakteristika takvih blogova je utjecaj na rezultate analize blogova. U čestom broju slučajeve generirani su algoritamski, pa su sustavi za analizu blogova u znatnoj mjeri sposobni prepoznati algoritamski stvorene nakupine teksta, koje u pozitivnom svjetlu predstavljaju određeni proizvod i izolirati ih iz uzorka za analizu. Takvi tekstovi osim što su primarno fokusirani na utjecanje ishoda rezultata analize, u sebi sadrže elemente koje na prvi pogled nalikuju na personalizirane tekstove kojima nije primarna funkcija promocija proizvoda.

Gdje smo tu mi ?

Hrvatska generalno, što se tiče primjene data mining metoda čak i nad strukturiranim podacima kaska za ostatkom svijeta. Vjerojatno će proći još mnogo vremena dok ova metodologija koja informacije crpi iz nestrukturiranih podataka zaživi unutar naših prostora. Pitanje koje se nameće samo po sebi proizlazi i iz činjenice problematike jezika. Kako smo mala nacija, ekvivalentno tome

proizvodimo količinski značajno manji broj blogova nego neka nacija s primjerice engleskog govornog područja.

Bilo kako bilo, kada jednom ova tehnologija stupi na naše prostore zasigurno će imati utjecaja na trendove istraživanja tržišta i unjeti novu dimenziju u metodologiju planiranja proizvoda i usluga. Prednosti ovakvog pristupa obradi tržišta su višestruke. Reduciraju troškove, selektiraju relevantan uzorak iz perspektive pronalaženja osoba koje svojevrijedno izražavaju stav o nečemu bez poticanja, što je dosta značajno za kvalitetu samog uzorka nad kojim se provodi istraživanje.

Ono što nam jedino preostaje jest zaron u beskonačne nizove tekstova, u nastojanju pronalaženja skrivenog blaga koje će u neka možda ne tako dalekoj budućnosti ugledati svjetlo dana.

