

Zašto je Euklid veći BI autoritet od Jamesa Bonda

Čest motiv koji je zaokupljao slavne matematičke umove, bilo je postojanje analitičkog instrumentarija u vidu matematičkih izraza, koji bi bili u stanju proračunati/predvidjeti prirodne pojave, dati položaje svih nebeskih tijela i predstaviti svijet u kojem živimo poput preciznog satnog mehanizma uz pomoć matematičkih jednažbi.

Među tim zanesenjacima bio je i slavni Pierre Simon Laplace, koji je u uvodu svoje knjige "Teorija vjerojatnosti", nastale davne 1812, zamišljao, kako on kaže, Inteligenciju koja bi u danom trenutku imala podatke o svim prirodnim silama, smještaju tvari od koje se sastoji svijet, i koja bi imala sposobnost matematičke analize svih tih podataka. Takva Inteligencija bi prema njegovu razmišljanju bila sposobna izgraditi formulu kojom bi se moglo obuhvatiti, kako gibanje najvećih svemirskih tijela, tako i gibanje atoma, te ništa ne bi bilo neizvjesno za tu vrstu Inteligencije, pa bi ona sa lakoćom mogla predviđati i budućnost.

Od ovog nadahnutog razmišljanja su prošla stoljeća, a mi i dalje živimo u svijetu kojeg najpreciznije možemo okarakterizirati Lorenzovim pitanjem: "Da li zamah leptirova krila u Brazilu može izazvati tornado u Teksasu?". Umjesto teorije koja bi se ukratko mogla usporediti sa satnim mehanizmom i koja bi bila najbliža Laplaceovim razmišljanjima, sve više posežemo u disciplinu teorije kaosa, kako bismo koliko toliko uspjeli objasniti svijet oko sebe. Mi u stvari živimo u svijetu u kojem vladaju pravila kaosa, u kojem je sve očitije da su Laplaceova razmišljanja, tek romantičarski izleti nadahnutog matematičkog uma.

No, da li ipak postoji barem nešto što bi bilo na tragu Inteligencije kakvu je zamišljao Laplace, da li barem postoji nešto između romantičarskih snova matematičkog uma i čistog kaosa što nam može pomoći, kako u percepciji svijeta u kojem živimo, tako i tržišne okoline u kojoj svi skupa obavljamo svoje svakodnevne poslovne aktivnosti?

Odgovor je *Business intelligence*.

Termin, koji je često puta u pogrešnim interpretacijama bio utjelovljenje duha Jamesa Bonda, u stvari ni izdaleka ne koketira sa brzim automobilima, atraktivnim špijunskim scenama, tajnim oružjima, borilačkim vještinama i ljepoticama sa naslovnih stranica. U filmu o BI-u neki su od glavnih junaka živjeli, za Bondovske uvjete vrlo povučenim i ujednačenim životom još u staroj Grčkoj. Pogađate, riječ je o matematičarima poput Euklida i Pitagore. Oni se naravno nisu bavili BI-om kao disciplinom, ali spadaju među odabrano društvo zaslužno za otkrića koja su sastavni dio suvremenih metoda za otkrivanje znanja iz velikih količina podataka. Članovi tog društva su još Bayes, Pascal, Laplace, Newton, Lobachevski, Gauss, te mnogi koji su svojim otkrićima omogućili razvoj softvera baziran na suvremenim metodama za otkrivanje znanja kakvog poznajemo danas. Primjerice, otkrića jednog Euklida sastavni su dio algoritama za klasteriranje, posebne kategorije neuronskih mreža (samoorganizirajuće mape), te čitavog niza varijacija algoritama za otkrivanje sličnosti. Euklid nije usamljen slučaj, jer gotovo da i ne postoji algoritam za otkrivanje znanja koji ne sintetizira niz otkrića od stare Grčke pa do današnjih dana. Blagodati razdoblja u kojem živimo očituje se i u tome, što u matematičkim modelima za strojno učenje možemo iskoristiti sva ova otkrića, pretočiti ih u algoritmizirane postupke s ciljem obrade nepregledne količine informacija u potrazi za znanjem unutar naših poslovnih podataka.

U svijetu Jamesa Bonda, podaci su oskudan resurs, saznaju se uglavnom uz škripu automobilskih guma i pucnjeve iz pištolja.

Naše doba karakterizira drugi tip problema. Podaci nisu oskudni resurs, te se pojavljuje problem kako iz tog obilja podataka raspoloživih iz internih i eksternih resursa iznjedruti pravu informaciju koja će nam pomoći u svakodnevnom poslovanju. Umjesto škrife automobilskih guma i topline revolverске cijevi, BI živi u okruženju vrućine mikroprocesora zagrijanog od proračuna neuronskih mreža, stabala odlučivanja, Bayesovih mreža, Fuzzy logike,

Samorganizirajućih mapa, klasteriranja i asocijativnih algoritama, oživljavajući duhove Euklida, Bayesa i Gaussa koji lebde između linija programskog koda.

Iako su algoritmi za strojno učenje nastali uglavnom u posljednjih nekoliko desetljeća i sintetiziraju čitav niz otkrića, svakako trebamo istaknuti ulogu znanstvenika koji su zaslužni za njihov razvoj, i koji su te algoritme razvijali imajući u vidu njihovu praktičnu primjenu.

Vratimo se Inteligenciji o kojoj je maštao Laplace. Danas je u svjetlu teorije kaosa više nego jasno, da takva savršena Inteligencija može egzistirati jedino kao plod mašte. No ne treba očajavati, jer uz pomoć inteligentnih algoritamskih metoda strojnog učenja i pretraživanja uzoraka danas se možemo pohvaliti da smo ipak sposobni pronaći pravilnosti i zakonitosti u sustavima koji se ponašaju kaotično i nepredvidivo. Ove inteligentne algoritamske metode spadaju u disciplinu koja se zove *data mining*, i sastavni su dio BI-a.

Metodama *data mininga* možemo precizno segmentirati tržište, otkriti profil tipičnog klijenta određene vrste proizvoda, njegove sklonosti ka kupnji i motivatore za *cross selling*. *Data mining* nam može pomoći u detektiranju potencijalnih prijevара na prodajnim mjestima, otkriti koji profil našeg klijenata je sklon prihvatiti ponudu konkurencije i zašto, predvidjeti trendove pojava, te otkriti sličnosti među određenim tržišnim kategorijama.

Ako primjerice želimo saznati, da li postoji stanovita pravilnost u trendovima prodaje u nekoj od trgovina (ako smo vlasnik lanaca trgovina) početkom tjedna i zašto, koji bi se proizvodi mogli kupovati u paru, kolika je vjerojatnost povećanja prodaje grupe proizvoda X ako im se u tom periodu spusti cijena za 10%, te da li bi se povećala prodaja artikla Y ako se kupuje u paru sa artiklom Z i za koliko posto, pomoć nam u ovoj vrsti analize može pružiti *data mining*.

Iako se ovaj zahtjev na prvi pogled čini vrlo kompleksnim, odgovore na sva ova pitanja moguće je dobiti obradom naših povijesnih transakcijskih podataka., te primjenom adekvatnih modela za strojno učenje.

Ovdje treba istaknuti činjenicu, da je u provođenju *data mining* analiza presudan faktor čovjek, te da softver sam po sebi ne rješava automatikom ovakve vrste problema. Ekspert ovdje ima presudnu ulogu, a softver mu služi tek kao pomoćni alat za konstrukciju i provedbu analitičkih modela.

Svjesni činjenice da nikada nećemo dostići ideal Laplaceovog sna o savršenoj Inteligenciji, ostaje nam jedino utjeha da će strojevi budućnosti imati puno veću moć po pitanju procesiranja i memoriranja podataka, što će zasigurno pogodovati osnaživanju i razvoju novih algoritama *data mininga* koji će na taj način postajati "puno pametniji". Usprkos svemu, jednu činjenicu je nemoguće osporiti, a to je da će čitav niz bivših, sadašnjih i budućih znanstvenih otkrića skladno surađivati u tom virtualnom prostoru u potrazi za znanjima duboko sakrivenim u nepreglednoj masi podataka .

© dr.sc. Goran Klepac
www.goranklepac.com
goran@goranklepac.com

Dr.sc. Goran Klepac, radi kao direktor Business intelligence direkcije u Raiffeisen Consultingu. Dugi niz godina radio je kao konzultant u domeni otkrivanja znanja iz baza podataka u različitim granama gospodarstva. Trenutno u svojstvu konzultanta vodi veći broj projekata vezanih uz potporu odlučivanju i otkrivanje znanja iz baza podataka.

Koautor je knjige *Poslovna inteligencija* i autor je knjige *Primjena inteligentnih računalnih metoda u managementu*, te suradnik na *Englesko-hrvatskom informatičkom enciklopedijskom rječniku*.

Predavač je na poslijediplomskom znanstvenom studiju *Informatički management* na Ekonomskom fakultetu u Zagrebu. Aktivni je sudionik niza znanstvenih i stručnih konferencija iz domene *data mininga*. Trenutno sa prof. dr.sc. Željkom Panianom radi na novoj knjizi radnog naslova "*Poslovna inteligencija kroz poslovne slučajeve*".