

Analiza veze osobina ličnosti i frekvencija slanja emotikona aproksimiranih Zipfovom raspodelom

Dunja Spasić, Kristina Popović

Matematički fakultet

10. septembar 2020.

Sažetak

Istraživanje univerziteta *Edge Hill University* o načinju korišćenja emotikona u onalajn razgovorima je pokazalo da postoji povezanost između načina slanja emotikona i nekih od pet glavnih osobina ličnosti. U ovom radu se formalizuje korišćenje emotikona tako što se frekvencije emotikona aproksimiraju Zipfovom raspodelom. Kolmogorov-Smirnov testom se određuje odgovarajuć koeficijent raspodele. Koeficijenti dobijenih raspodela se posmatraju kao jedno od obeležja, dok su ostala obeležja rezultati testa *Big Five*. Između koeficijenta Zipfove raspodele i svake od osobina ličnosti se izračunava korelacija. Nije dobijena nikakva značajna korelacija ni za jednu osobinu ličnosti.

Sadržaj

1	Uvod	2
2	Ideja rada	2
3	Metode	2
4	Rezultati	4
5	Diskusija i zaključak	4
	Literatura	5

1 Uvod

Razvitak i evolucija ljudske komunikacije je od uvek bila čvrsto povezana sa pisanim jezikom ali i jezikom slika. Slike su prvi zabeleženi vid pisanih informacija. Iako je tekstualni zapis danas mnogo precizniji i informativniji, u digitalnoj komunikaciji informativnost emotikona nije zanemarljiva.

Istraživanje sprovedeno na univerzitetu *Edge Hill University* [3] je pokazalo da ljudi na veoma različite načine tumače poruke poslate bez emotikona i iste te poruke poslate sa emotikonima. Emotikoni su uspeli da uvedu neverbalnu komunikaciju u digitalne razgovore.

Glavna ideja pomenutog istraživanja je bila da se utvrdi da li postoji povezanost sa načinom i frekvencijom slanja emotikona i nekim osobinama ličnosti. Istraživači su za test ličnosti koristili test *Velikih 5* [1], tj. pet osobina ličnosti koje su nezavisne jedne od drugih i iz kojih mogu da se izvedu mnoge druge osobine. Pet glavnih osobina su otvorenost za nova iskustva, savesnost, ekstrovertnost, dopadljivost i neurotizam. U istraživanju je grupi ispitanika dato da urade test ličnosti a onda su istraživači zajedno sa ispitanicima pogledali njihove fejsbuk profile. Zatim su slobodnom procenom doneli zaključak koje emotikone ispitanik najviše koristi i da li često koristi emotikone ili ne. Primetili su da postoji povezanost između frekvencije slanja emotikona i dve osobine ličnosti, ekstrovertnosti i otvorenosti ka novim iskustvima.

2 Ideja rada

Istraživanje Edž Hil Univerziteta je dalo interesantne rezultate ali može se primetiti da su ocene načina komuniciranja emotikonima dobijene neformalno, ličnom procenom istraživača. Ideja našeg rada je da formalizuje ove informacije o slanju emotikona, tj. da se svakoj osobi dodeli neka ocena koja bi opisivala način na koji ona komunicira emotikonima.

Ako emotikone posmatramo kao reči u komunikaciji, moglo bi se pretpostaviti da će oni onda imati i istu raspodelu kao reči u razgovoru. Zanimljivo je da u svakom jeziku sveta, reči koje se koriste imaju istu raspodelu. Kada bi se prebrojale reči iskorišćene u nekoj knjizi ili razgovoru i kada bi se one sortirale i predstavile štapićastim dijagramom, dobijena raspodela bi bila Zipfova raspodela.

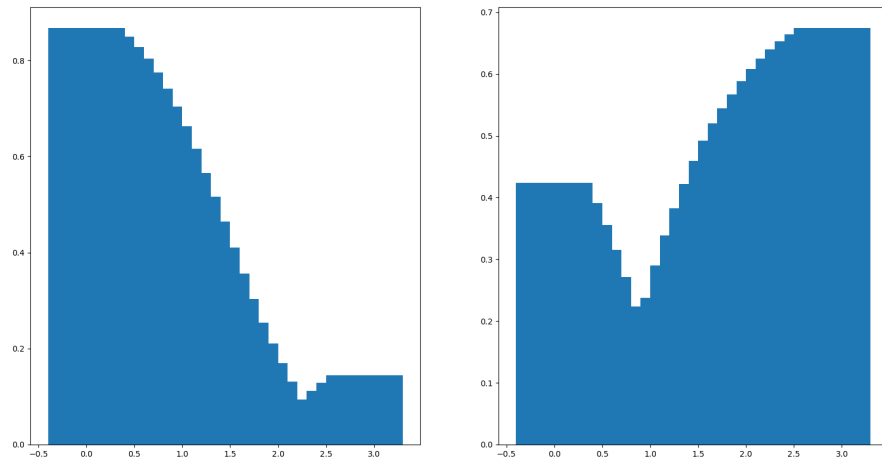
Zipfova raspodela je aproksimacija Zipfovog zakona. Zipfov zakon je empirijski zakon koji kaže da su frekvencije korišćenih reči u tekstu obrnuto proporcionalne njihovom rangu u tabeli frekvencija svih reči. [2] Ovo znači da će se prva najkorišćenija reč pojaviti duplo više puta od druge, tri puta više od treće načešće reči itd.

Stoga, u ovom radu je pokušano da se emotikoni koje svaka osoba šalje u razgovorima sortiraju opadajuće po frekvenciji, a zatim normalizuju, a zatim bi im se dodelio koeficijent Zipfove raspodele koja najvernije predstavlja frekvencije emotikona. Koeficijent Zipove raspodele bi onda bio formalna numerička ocena načina slanja emotikona, koja bi kasnije mogla da se uporedi sa rezultatima testa *Velikih pet* (referenca sajta).

3 Metode

Kao uzorak su uzeti razgovori sa aplikacija *WhatsApp*, *Messenger* i *Telegram*. Podaci su obrađeni tako da se iz razgovora, za svakog člana tog

razgovora izdvaja mapa koja mapira frekvenciju korišćenja svakog emotikona na *UNICODE* kod tog emotikona. Zatim se tako napravljena mapa sortira. Dobijeni podaci o emotikonima se onda Kolmogorov-Smirnovim testom upoređuju sa Zipfovom raspodelom za različite koeficijente. Na slici 1 su prikazani rezultati Kolmogorov-Smirnov testa za svaki testiran koeficijent Zipfove raspodele, sa korakom 0,1.



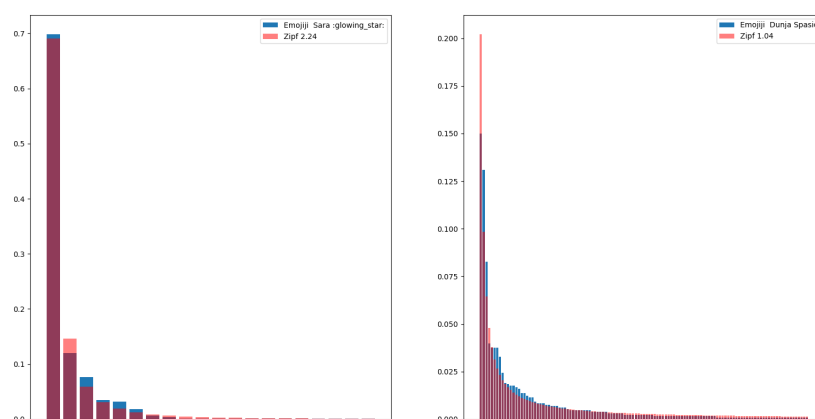
Slika 1: Rangovi aproksimacije (y-osa) za svaki testiran koeficijent Zipfove raspodele (x-osa).

Za najmanji rezultat se proverava značajnost testa, i izbacuju se iz uzorka sve osobe čija vrednost test statistike upada u kritičnu oblast. Za svaki razgovor se računaju koeficijenti raspodele i zatim se za sve sakupljene vrednosti koeficijenta raspodele određene osobe računaju srednja vrednost i disperzija. Za podatke se kasnije računa korelacija sa rezultatima testa osobina ličnosti.

4 Rezultati

Sakupljeni su podaci od 32 osobe, s tim što su neke osobe bile u više od jednog razgovora pa su za njih izračunate srednja vrednost i disperzija koeficijenata Zipfovih raspodela svakog razgovora. Za ostale je uzeta samo jedna zabeležena vrednost koeficijenta.

Na slici 2 su prikazane preklopljene raspodele emotikona i njoj najbliže Zipfove raspodele jednog od razgovora. Može se primetiti da prva osoba šalje značno manje emotikona, pa zato njoj odgovara veći koeficijent Zipfove raspodele.



Slika 2: Raspodela poslatih emotikona jednog od razgovora i njoj najbliža Zipfova raspodela.

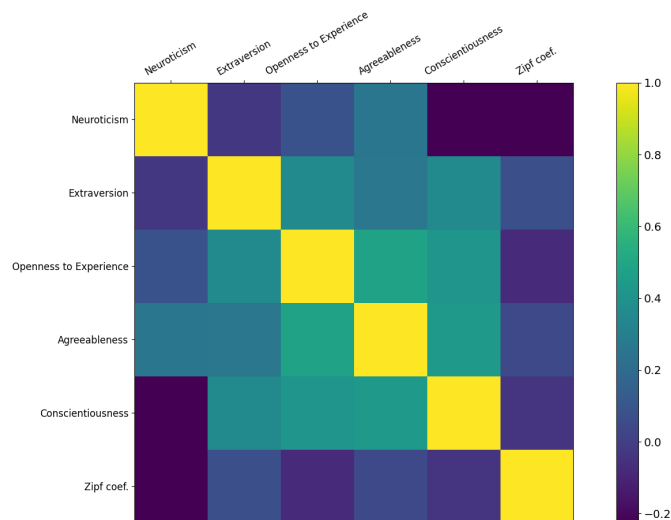
Za korelaciju između obeležja *Koeficijent Zipfove raspodele* i svakog drugog zabeleženog obeležja, tj. svake od pet gavlkih osobina ličnosti se dobijaju koeficijenti korelacije -0.22573 za koeficijent Zipfove raspodele i neurotizam, 0.071552 za koeficijent Zipfove raspodele i otvorenosti ka novim iskustvima, 0.045021 za koeficijent Zipfove raspodele i dopadljivost i -0.034568 za koeficijent Zipfove raspodele i savesnost. Matrica koeficijentata korelacija svih posmatranih obeležja je prikazana na slici.

Očigledno, ni jedana merena osobina ličnosti nema dovoljno značajnu korelaciju sa koeficijentom Zipfove raspodele. Po apsolutnoj vrednosti, najveću korelaciju sa Zipfovom koeficijentom ima neurotizam, -0.22573. Ova korelacija nije dovoljno velika da bi se smatrala značajnom.

5 Diskusija i zaključak

Iako je istraživanje Edž Hil univerziteta ukazalo na povezanost načina slanja emotikona i nekih osobina ličnosti, na osnovu formalizacije podataka o emotikonima opisanim u radu, nije pronađena dovoljno jaka korelacija među ove dve pojave.

Ovo ne mora nužno da znači da je istraživanje Edž Hil univerziteta došlo do netačnih rezultata. Treba uzeti u obzir, pre svega, da ova nu-



Slika 3: Matrica koeficijenata korelacije posmatranih obeležja

merička interpretacija ne uzima u obzir tip emotikona koji osobe najviše koriste, već samo njihovu raspodelu. Druga potencijalna razlika može da bude što je možda uzorak prvobitnog istraživanja bio veći od 32, pa možda na ovako malom uzorku nije mogao da se opazi obrazac.

Takođe, ovo ne znači da u potpunosti ne postoji korelacija između osobina ličnosti i koeficijenta Zipfove raspodele. Još jedna zanimljiva hipoteza za testiranje bi bila da li postoji korelacija između disperzije Zipfovog koeficijenta jedne osobe i dopadljivosti, jer dopadljivost podrazumeva delom i prilagođavanje sagovorniku, što bi moglo da ukaže i na prilagođavanje frekvencije slanja emotikona i tipa emotikona koji se šalje. Nažalost, u sakupljenom uzorku nije bilo dovoljno ispitanika kojima je bila zabeležena disperzija, pa ne bi imalo svrhe na ovom uzorku testirati tu hipotezu.

Literatura

- [1] Zvanična internet adresa testa *Big Five*. <https://bigfive-test.com/>.
- [2] Ramon Ferrer i Cancho and Ricard V. Solé. Two regimes in the frequency of words and the origins of complex lexicons: Zipf's law revisited. *Journal of Quantitative Linguistics*, 8(3):165–173, 2001.
- [3] Dr Linda Kaye. Zvanična internet adresa projekta *Multi-modal Communication*, year = 2019, note = <https://www.edgehill.ac.uk/psychology/research/research-groups/multi-modal-communication/>.