

Razvijanje hibridnog sistema koji koristi fazi logiku i duboko učenje za analizu osećanja

Aleksa Veselić

April 2023

Sažetak

Cilj ovog rada je da prikaže integraciju transformer modela sa fazi pravilima koja određuju vrednost čitanja komentara na Reddit-u na osnovu dužine, subjektivnosti i sentimenta. Želim da da pokažem efikasnost ovog pristupa u identifikaciji komentara koji su vredni vremena i truda za čitanje.

1 Uvod

Obrada prirodnog jezika (NLP) je područje veštačke inteligencije koje se bavi interakcijom između računara i ljudskih jezika. Glavni cilj NLP-a je da omogući računarima da razumeju, interpretiraju i generišu ljudske jezike. Veliki proboj u ovom polju nastaje objavljivanjem rada "Attention Is All You Need" [3] i uvođenjem nove arhitekture Transformer koja se pokazala da ima bolje performanse u poredjenju sa njenim prethodnicima kao što su RNN (rekurentne neuronske mreže), TF-IDF (učestalost termina - inverzna učestalost dokumenata), LSTM (dugotrajna kratkoročna memorija).

1.1 Problem

Društvena mreža Reddit ima gornje ograničenje dužine komentara 40 000. Usled toga navigacija kroz komentare i nalaženje komentara koji su relevantni za objavu ispod koje se nalaze može oduzimati mnogo vremena.

2 Model

Opisani model predstavlja nekoliko uzastopnih povezanih slojeva koji će dalje biti redom objašnjeni. Izlaz jednog sloja predstavlja ulaz narednog.

2.1 Embedding

Prvi korak jeste konvertovanje stringovskog ulaza u numeričku reprezentaciju. Ovde je korišćenja prekompilirana biblioteka Glove[2]. Prednost njenog korišćenja jeste smanjivanje vremena učenja modela. Neke od prednosti predstavljanja reči kao vektora u prostoru jesu sledeće:

Najbliži susedi Vektori koji se nalaze blizu jedan drugoga (u pogledu Euklidskog ili kosinusnog rastojanja) imaju slično lingvističko ili semantičko značenje. Sledeći primer ilustruje koje reči su susedi reči 'frog':

1. frog
2. frogs
3. toad
4. litoria
5. leptodactylidae
6. rana
7. lizard
8. eleutherodactylus

Linearne podstrukture: metrike sličnosti koje se koriste za evaluacije najbližih suseda proizvode jedan skalar koji kvantifikuje srodnost dve reči. Ova jednostavnost može biti problematična jer dve date reči skoro uvek pokazuju zamršenije odnose nego što se mogu obuhvatiti jednim brojem. Na primer, muškarac se može smatrati sličnim ženi po tome što obe reči opisuju ljudska bića; s druge strane, ove dve reči se često smatraju suprotnostima jer ističu primarnu osu duž koje se ljudi razlikuju jedni od drugih.

Da bi se na kvantitativan način uhvatila nijansa neophodna za razlikovanje muškarca od žene, potrebno je da model paru reči poveže više od jednog broja. Prirodan i jednostavan kandidat za uvećani skup diskriminativnih brojeva je vektorska razlika između dva vektora reči. GloVe je dizajniran da takve vektorske razlike obuhvate što je više moguće značenje određeno poredjenjem dve reči.

2.2 Multi-Head Attention

Da bi razumeli mehanizam rada višeglave pažnje (MHA) prvo ćemo uvesti mehanizam samo-pažnje (self-attention) koji omogućava Transformer modelu da računa pažnju (težinske koeficijente) za svaki par reči u ulaznom tekstu. Ovi koeficijenti pažnje se zatim koriste za ponderisanje reprezentacija reči u svakoj sekvenci, pri čemu se više pažnje poklanja relevantnijim rečima. Ovaj proces se izvodi iterativno za sve sekvence u tekstu, omogućavajući modelu da efikasno "sazna" kako različite reči međusobno zavise.

Mehanizam samo-pažnje se oslanja na skalarni proizvod (dot product) pažnje izmedju svih parova reči u rečenici (queries). Ovo omogućava modelu da nauči kompleksne interakcije izmedju reči i identifikuje njihove medjusobne odnose. Takodje, mehanizam samo-pažnje je paralelan, što ga čini efikasnim za obradu velikih sekvencijalnih podataka.

$$Attention(Q) = softmax(\frac{QQ^T}{\sqrt{d_Q}})Q$$

Analogno ovome, mehanizam višeglave pažnje ulazne vektore deli na N delova (N je broj glava), na njima računa samo-pažnju nakon čega ih spaja u matricu odgovarajuće veličine. Ovo nam omogućava da dublje razbijemo reči na njihove disjunktne osobine uz pomoću kojih ćemo razaznati još veću semantičku sličnost dveju reči.

$$MultiHead(Q, N) = Concat(head_1, head_2, ..., head_N)$$

$$where head_i = Attention(Q[i * d_Q / N : (i + 1) * d_Q / N])$$

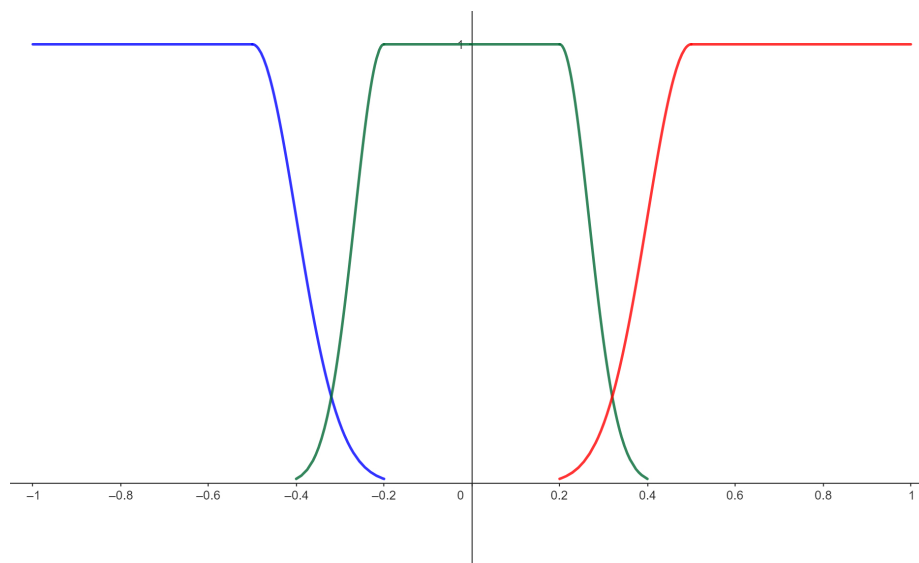
2.3 Feed-Forward Network

Mreža sa propagacijom unapred (FFN) obavlja nelinearne transformacije nad reprezentacijama reči kako bi naučio kompleksne oblike zavisnosti i interakcija medju rečima. FFN služi kao lokalni kontekstualizator, gde se za svaku reč izračunava posebna transformacija. Ovo omogućava modelu da nauči bogatije reprezentacije reči koje su prilagodjene kontekstu svake reči u tekstu.

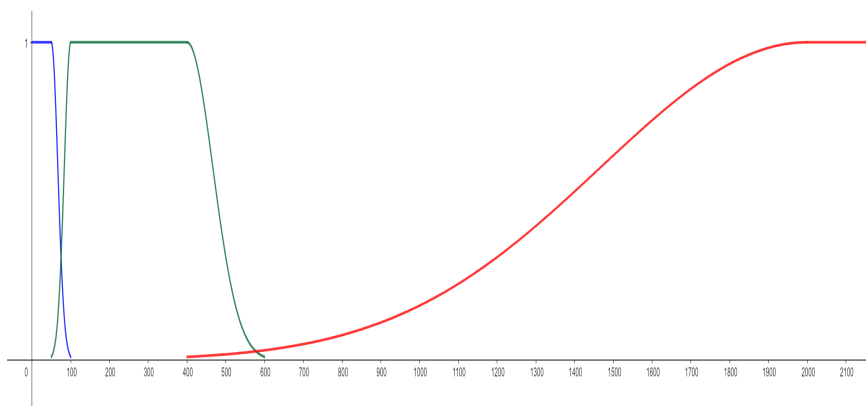
2.4 Fazi pravila

Pomoću fazi pravila želimo da unesemo domensko znanje prilikom rešavanja konkretnog problema. Dimenzije koje svaki tekst poseduje su sledeće:

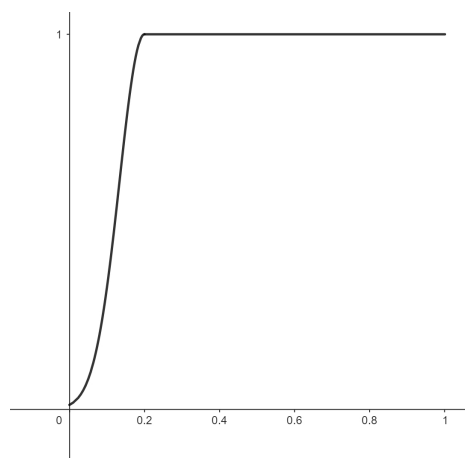
- **Sentiment(osećanje)** - izlaz iz neuronske mreže može biti -1 (negativan tekst), 0 (neutralan) i 1 (pozitivan)
- **Dužina** - broj karaktera u tekstu
- **Subjektivnost** - koristeći *leksikon subjektivnosti*[4] računamo relativnu ocenu gde je 0 objektivno a 1 subjektivno napisan tekst



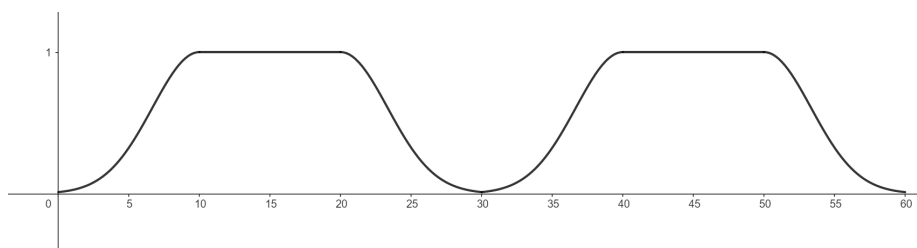
Redom: negativni, neutralni i pozitivni sentiment.



Redom: kratki, srednji i dugi tekstovi



Relativna subjektivnost teksta.



Redom: mišljenja, eseji.

Definišemo 3 nova tipa teksta na osnovu pravila:

- **Mišljenja**

$(srednje\ dugi) \wedge (subjektivni) \wedge (pozitivni)$

$(srednje\ dugi) \wedge (subjektivni) \wedge (negativni)$

- **Eseji**

$(dugi) \wedge (objektivni) \wedge (pozitivni)$

$(dugi) \wedge (objektivni) \wedge (neutralni)$

$(srednje\ dugi) \wedge (objektivni) \wedge (pozitivni)$

$(srednje\ dugi) \wedge (objektivni) \wedge (neutralni)$

- **Beskorisni** - ostalo

3 Rezultati

Prikazani su rezultati rada Transformer modela na skupu 'Twitter and Reddit Sentimental analysis Dataset[1]'. Korišćeni su samo komentari na Reddit društvenoj mreži zato što, za razliku od komentarana Twitter-u, nisu ograničeni malim brojem (280) karaktera. Dužina samih komentara nam predstavlja još jednu dimenziju u podacima.

Trening instance	Slojevi FFN	Epohe	Broj glava	Train %	Test %
10k	4	1	10	30.88	30.66
15k	4	1	10	36.76	36.52
20k	4	1	10	42.26	42.19
25k	4	1	10	43.16	43.17
30k	4	1	10	43.24	43.36
37k	3	1	5	30.56	30.73
37k	3	2	5	44.44	44.27
37k	3	2	5	44.44	44.27
37k	3	3	5	44.38	44.43
37k	3	4	5	44.34	44.56
37k	3	1	10	42.25	42.48
37k	3	2	10	43.21	43.17
37k	4	1	5	46.74	46.65
37k	4	2	5	42.70	42.70
37k	4	3	5	42.82	42.98
37	4	3.125	5	42.79	42.82
37	4	3.25	5	42.72	42.72
37k	4	1	10	44.31	44.14
37k	4	10	10	54.92	54.84
37k	4 ¹	1	10	54.91	54.84
37k	5	1	10	42.70	42.70
37k	5	2	10	42.70	42.70
37k	4	1	15	56.79	57.55

4 Zaključak

Transformer arhitektura pokazuje sjajne rezultate i otvorena je na mogućnost paralelnog izračunavanja. Dalje unapredjenje projekta bi se sastojalo iz kompleksnije semantičke analize teksta kao na primer: različito značenje pojedinih reču u odnosu na to koju ulogu imaju u rečenici.

Takodje, model sa većim brojem parametara koji će iza sebe imati moćniji hardver na kome će biti treniran. Dodatno možemo koristiti poziciono enkodiranje koje može biti funkcija čije parametre uči neuronska mreža ili kao predefinisani skup pravila.

¹Više neurona nego prethodni model sa istim brojem slojeva

References

- [1] Chaithanya Kumar A. Twitter and reddit sentimental analysis dataset.
- [2] Jeffrey Pennington, Richard Socher, and Christopher D. Manning. Glove: Global vectors for word representation. In *Empirical Methods in Natural Language Processing (EMNLP)*, pages 1532–1543, 2014.
- [3] Ashish Vaswani, Noam Shazeer, Niki Parmar, Jakob Uszkoreit, Llion Jones, Aidan N Gomez, Łukasz Kaiser, and Illia Polosukhin. Attention is all you need. In *Advances in Neural Information Processing Systems*, pages 5998–6008, 2017.
- [4] Theresa Wilson, Janyce Wiebe, and Paul Hoffmann. Recognizing contextual polarity in phrase-level sentiment analysis. In *Proceedings of Human Language Technology Conference and Conference on Empirical Methods in Natural Language Processing*, pages 347–354, Vancouver, British Columbia, Canada, October 2005. Association for Computational Linguistics.