



УНИВЕРЗИТЕТ У НОВОМ САДУ
ПРИРОДНО-МАТЕМАТИЧКИ ФАКУЛТЕТ
ДЕПАРТМАН ЗА МАТЕМАТИКУ И
ИНФОРМАТИКУ



Key Features Net i njene primene

Seminarski rad

Miloš Vujašinić

Novi Sad, jul 2020

Sadržaj

Sadržaj	1
Uvod	2
1 Poglavlje 1	3
Zaključak	4
Literatura	5

Uvod

Autoenkodori su već godinama zlatni standard u smanjenju dimenzionalnosti podataka. Način na koji rade se pokazao kao veoma efikasan u otklanjanju šuma i dopunjavanju podataka koji su oštećeni. Motivisan datim primerima, ovaj rad pokušava da prikaže novi način gledanja na smanjene dimenzionalnosti podataka mašinskim učenjem. Zatim, primenom iznesenih ideja i nekih od principa koji se nalaze u osnovi autoenkodera se uvodi model neuronske mreže koji za cilj da iz podataka koji se prosleđuju modelu izdvoji najbitnije odlike za klasifikaciju. Kroz ovaj postupak se takođe razmatraju novi načini treniranja i evaluacije modela, a na kraju se rezultati datog modela porede sa rezultatima tradicionalnih autoenkodera.

1 Poglavlje 1

Autoenkoderi su neuronske mreže koje se treniraju da oponašaju identičko preslikavanje, to jest da ulaz preslikava na samog sebe. Ovo samo po sebi neizgleda veoma korisno, međutim autoenkoderi, pored navedenih, sadrže i osobinu da je broj neurona u najmanjem skrivenom sloju mreže manji od broja neurona ulaznog i izlaznog sloja. Ovo za implikaciju ima da se prilikom prolaska podataka kroz mrežu u datom sloju nalazi reprezentacija podataka za čije je predstavljanje potreban manji broj memorijskih jedinica od originalnih podataka i ovo može da varira od nekoliko jedinica, pa do redova veličina manje. Dati skriveni sloj se naziva *usko grlo* (engl. *bottle neck*) i on deli neuronsku mrežu na dva dela: od ulaza do sebe i od sebe do izlaza. Ovi delovi se redom nazivaju *enkoder* i *dekoder*, a oni se koriste za preslikavanje originalnih podataka u reprezentaciju smanjenih dimenzija i nazad.

Prilikom treniranja autoenkodera se greška računa kao greška između svakog neurona ulaza i njemu odgovarajućeg neurona izlaza. Ovo je veoma intuitivno rešenje ako želimo da rezultati na izlazu budu naizgled što sličniji ulaznim podacima. Međutim dato preslikavanje je u praksi gotovo nemoguće napraviti tako da bude savršeno. Kao posledica dolazi do izvesne greške između originalnih podataka i rezultata mreže koji možda neizgledaju značajno ili su nam gotovo neprimetni, ali, teoretski, mogu značajno da utiču na rezultate klasifikacije ako bi se rezultati autoenkodera pustili kroz klasifikator i uporedili sa rezultatima klasifikacije originalnih podataka. Na primer, ako imamo klasifikator koji traga za odlikom koja je sadržana u veoma malom broju memorijskih jedinica u poređenju sa veličinom jedne instance podataka, a autoenkoder zaključi da bi uvrštavanje date odlike samo povećalo grešku, jer bi njenim uvrštavanjem bilo onemogućeno uvrštavanje neke druge odlike koja je iz pogleda greške autoenkodera važnija. Kao posledica, redukcija podataka ovim postupkom je ne potencijalno samo beskorisna, nego i čini podatke beskorisnim ako želimo da ih klasifikujemo ili, još gore, koristimo za treniranje klasifikatora.

Diskutovani problem svakako postavlja pitanje kako se može prevazići, ali pre davanja odgovora na njega, potrebno je da definišemo podelu koja će da omogući uočavanje ovog problema jasnijim i lakšim. Iz prikazanog rezonovanja se jasno može uvideti da se autoenkoderi mogu trenirati ili u svrhu očuvanja izgleda podataka ili u svrhu očuvanja onoga što podaci ustvari jesu, odnosno njihove suštine. Treba imati u vidu da podaci imaju gotovo bezbroj različitih suština u zavisnosti od toga iz kog se ugla posmatraju, to jest da se u zavisnosti od pitanja koje se postavlja o njima njihova suština menja i odatle možemo da spojimo ideju o suštini sa problemom klasifikacije tako što ćemo reći da pitanje koje postavlja klasifikacioni problem određuje suštinu podataka o kojoj u datom trenutku pričamo. Ove ideje se koriste za podelu definisanu u definiciji 1.

Definicija 1 (Podela autoenkodera po načinu treniranja) *Autoenkoder koji za čije se treniranje koristi znanje o samim podacima se naziva izgledni autoenkoder, dok autoenkoder za čije se treniranje koristi znanje o klasama kojima podaci pripadaju se naziva suštinski autoenkoder. Treba imati u vidu da date osobine nisu disjunktne, pa postoji i izgledno-suštinski autoenkoder.*

Zaključak

Literatura

- [1] Pierre Baldi. Autoencoders, unsupervised learning, and deep architectures. In Isabelle Guyon, Gideon Dror, Vincent Lemaire, Graham Taylor, and Daniel Silver, editors, *Proceedings of ICML Workshop on Unsupervised and Transfer Learning*, volume 27 of *Proceedings of Machine Learning Research*, pages 37–49, Bellevue, Washington, USA, 02 Jul 2012. PMLR.