

Klasifikacija zamišljenih reči

Lea Kojičić mi22365

9. mart 2024.

Sadržaj

1	Uvod	2
2	Skup podataka	2
2.1	Pretprocesiranje	2
3	Rešenje problema	2
3.1	Klasifikacija Slučajnom šumom	2
3.1.1	Parametri modela	3
3.1.2	Ugnježdjena unakrsna validacija	3
3.2	Implementirani modeli	3
4	Zaključak	4

1 Uvod

Problem klasifikacije zamišljenog govora se odnosi na klasifikaciju reči koje neki subjekat zamisli. Ključni podaci su dakle misli, koje samo predstavljaju moždanu aktivnost u određenim regionima mozga. Ovi podaci se mogu sakupiti invazivnim ili neinvazivnim metodama; pri čemu prve daju podatke koji rezultuju preciznom klasifikacijom, ali nameće mnoštvo etičkih problema, te nije pogodan za širu upotrebu. Moj projekat klasifikuje reči na osnovu neinvazivno prikupljenih podataka - EEG signala. Koristila sam 3 modela mašinskog učenja, trenirana na podacima o ukupno 2 subjekta. Kako je problem klasifikacije neinvazivno prikupljenih signala i dalje veoma težak, rezultati mog rada su takođe daleko od zadovoljavajućih. Dalji naponi u ovoj oblasti istraživanja treba da ukažu da li su razlog tome loše odabrani modeli, pretprocesiranje, nepreciznost neinvazivne tehnologije ili nešto četvrto.

2 Skup podataka

Korišćeni podaci [NPR⁺22] predstavljaju signale snimljene zamišljanjem jedne od četiri reči: gore, dole, levo ili desno. Iako su prikupljeni za 10 učesnika, u projektu su (zbog memorijskog ograničenja) korišćeni podaci o samo 2 učesnika. Prikupljanje podataka o učesnicima odvijalo se u 3 sesije, gde svaka ima 100 ponavljanja. Jedno ponavljanje podrazumeva sledeće:

1. Ponavljanje počinje u trenutku $t=0$, kada se učesnik umiruje i priprema za stimulus
2. U intervalu $t \in [0.5s, 1s]$ česniku se prikazuje strelica na ekranu
3. Naredne 2.5s su označene kao period *akcije* kada učesnik zamišlja reč koja odgovara prikazanom smeru

2.1 Pretprocesiranje

Kako su vremena izvršavanja akcije (zamišljanja reči) fiksirana, prvo je bilo potrebno izdvojiti signale iz vremenskog intervala $[1, 3]$. Ovo je učinjeno metodom *Select time window*, koja po preporuci izvornog rada (citat) pomoću frekvencije od 256Hz preslikava vremenski segment u odgovarajući segment signala.

Naredni korak je izdvajanje samo određenih frekvencija $\alpha, \beta, \gamma, \delta$ i θ moždanih talasa, koje se smatraju relevantnim za rešavanje datog problema. Korišćene su frekvencije preporučene u radu [PDSC⁺22].

Poslednji korak u fazi pretprocesiranja je izdvajanje *karakterističnih prostornih obrazaca* moždane aktivnosti. Ova tehnika (eng. Common spatial patterns) je česta u problemima klasifikacije EEG signala, stoga je dostupna u Pajton MNE biblioteci. EEG signal je predstavljen kao trodimenzionalni tenzor dimenzija (broj uzoraka, broj elektroda, broj vremenskih tačaka). CSP model pronalazi prostorne filtere koji maksimalno razlikuju različite klase u trening skupu, i na osnovu toga transformiše podatke. Ova transformacija nad podacima smanjuje dimenzionalnost i omogućava lakše treniranje modela. Međutim, jasno je da nepažljiva primena CSP transformacije značajno može ugroziti rezultate klasifikacije.

3 Rešenje problema

Klasifikacija je vršena korišćenjem bibliotečkog (sklearn) modela Slučajne Šume, kao i implementiranim modelima 2d Konvolutivne mreže i potpuno povezane Neuronske mreže.

3.1 Klasifikacija Slučajnom šumom

Slučajne šume su agregirani model stabla odlučivanja - ansambl. To znači da su sastavljene od više stabla odlučivanja čije su greške klasifikacije nezavisne. Ova stabla učestvuju u klasifikaciji slučajne šume davanjem glasa", odnosno, predviđanje slučajne šume je najbornoje predviđanje pojedinačnih stabla odlučivanja. Maksimalna ostvarena tačnost predviđanja upotrebom slučajne šume je 0.27 (nešto iznad šanse).

Za klasifikaciju je korišćena Sklearn biblioteka implementacija Slučajne šume (`sklearn.ensemble.RandomForestClassifier`) a za pretragu hiperparametara - Ugnježđena unakrsna validacija (takode biblioteka). Navedeni hiperparametri se odnose na stabla odlučivanja i kontrolišu preprilagođavanje modela.

3.1.1 Parametri modela

- `n_estimators`
broj stabla odlučivanja koja učestvuju u ansamblu
- `max_depth`
maksimalna dubina pojedinačnog stabla odlučivanja
- `min_samples_split`
minimalan broj podataka koji mora da se nađe u čvoru pojedinačnog stabla
- `min_samples_leaf`
minimalan broj podataka koji mora da se nađe u listu pojedinačnog stabla. Podela na višim nivoima će se izvršavati uzimajući u obzir ovaj parametar.

3.1.2 Ugnježđena unakrsna validacija

Ugnježđena unakrsna validacija se vrši pomoću spoljašnje i unutrašnje petlje.

U spoljašnjoj petlji se kompletan skup podataka deli na test skup i skup koji će unutrašnjoj petlji služiti za trening i validaciju (train-validation skup). Nakon pretrage hiperparametara, i na osnovu ocene na validacionom skupu (u okviru unutrašnje petlje) - bira se najbolji model koji se testira na test skupu. Ovaj proces se ponavlja 3 puta.

Unutrašnja petlja "kreira" modele na osnovu svih kombinacija parametara prosleđenih u `params-grid`. Tako kreirane modele trenira na i validira metodom unakrsne validacije. Validacioni rezultati se pamte i služe za izbor konačnog modela na kraju spoljašnje petlje. Sve ovo je omogućeno klasom `GridSearchCV`, u Sklearn biblioteci.

3.2 Implementirani modeli

Pri klasifikaciji Konvolutivnom mrežom nisam koristila CSP transformaciju u pretprocesiranju. Ovo nije ni pomoglo ni odmoglo klasifikaciji, obzirom da su i najbolja i prosečna tačnost modela jednaka šansi - 0.25. Iako je za treniranje potpuno povezane neuronske mreže korišćena CSP transformacija, rezultati su slični - maksimalna tačnost modela je 0.26 a prosečna 0.25.

Pri odabiru arhitektura vodila sam se hardverskim ograničenjima, 2d Konvolutivnim mrežama koje sam nalazila na internetu, i konačno kreirala arhitekturu na osnovu generalnog utiska i zapamćenih arhitektura.

Arhitektura Konvolutivne mreže je sledeća:

- 3 konvolutivna sloja sa jezgrom veličine 3x3 i ReLU aktivacionom funkcijom
- Prva 2 konvolutivna sloja dodatno smanjuju dimenziju MaxPool metodom, dok treći konvolutivni sloj koristi Dropout sa verovatnoćom 0.3
- Poslednji sloj mreže je potpuno povezani sloj sa Dropout-om verovatnoće 0.2

Potpuno povezana mreža se sastoji iz 3 linearna sloja, i korišćena je ReLU aktivaciona funkcija.

4 Zaključak

Moji rezultati, kao i većine radova [LBBPM22] koji klasifikuju EEG signale zamišljenog govora su veoma loši. Uticaj na moje rezultate svakako imaju i memorijska (podaci za samo 2 učesnika) ograničenja, proizišla iz hardverskih. Ono što se u praksi pokazalo, a u mom radu se nazire - jeste da je CSP transformacija značajna za bolju klasifikaciju. Ostavljen je prostor za pokušaje klasifikacije LSTM modelom, 1d Konvolucijom ili upotrebom Transformer-a. Razlog generalno lošim rezultatima sigurno leži i u malom broju podataka (korišćeni skup podataka je jedini javno dostupan), i nepreciznim mernim instrumentima. EEG tehnologija je sama po sebi prostorno veoma neprecizna, a dodatnu prepreku mogu predstavljati i jeftini uređaji korišćeni za merenje signala. Međutim, upravo instrumenti za merenje EEG signala su lako prenosivi, stoga pogodni za širu upotrebu i dalju primenu u Brain-Computer Interface-ima. Ostaje da se nadamo da je problem u neistraženosti problema, lošim računarskim pristupima - a ne u samom hardveru pri skupljanju podataka. Ja verujem da će ovaj problem uskoro biti rešen.

Literatura

- [LBBPM22] D. Lopez-Bernal, D. Balderas, P. Ponce, and A. Molina. A state-of-the-art review of eeg-based imagined speech decoding. *Front Hum Neurosci*, 16:867281, Apr 2022.
- [NPR⁺22] N. Nieto, V. Peterson, H. L. Rufiner, et al. Thinking out loud, an open-access eeg-based BCI dataset for inner speech recognition. *Sci Data*, 9:52, 2022.
- [PDSC⁺22] T. Proix, J. Delgado Saa, A. Christen, et al. Imagined speech can be decoded from low- and cross-frequency intracranial EEG features. *Nat Commun*, 13:48, 2022.