

Eksperyment 2

Hipoteza

Sieć rekurencyjna powinna sobie poradzić z dużą ilością cech, które mają charakter sekwencyjny, tzn. kolejne cechy przedstawiają zmianę wartości w czasie. Tak stworzony klasyfikator powinien uzyskać wyniki lepsze niż te przedstawione w oryginalnym artykule.

Sposób trenowania

Do trenowania wykorzystaliśmy rekurencyjną (LSTM) sieć głęboką. Dane uczące zostały podzielone na dwa zbiory - dane EEG oraz dane z urządzeń peryferyjnych. Następnie dane uczące były wykorzystywane do nauki każdej emocji oddzielnie.

Zaimplementowana sieć składa się z 20 neuronów LSTM oraz jednego neuronu wyjściowego, która odpowiada na pytanie czy uczona emocja została wykryta.

Trenowanie następuje w batchach o wielkości 120 próbek.

Sieć została zaimplementowana w Kerasie i trenowana przy użyciu GPU.

Sposób testowania

Testowane jest 6 klasyfikatorów (po jednym dla każdej pary zbiorów uczący - uczucie). Każdy klasyfikator jest testowany za pomocą walidacji skrośnej z podziałem na 5 podzbiorów. Każdy podzbiór pełni raz rolę zbioru testującego, a reszta zbiorów pełni w tym czasie rolę zbiorów trenujących. Dla każdego takiego etapu wyliczamy accuracy i F1-score. Następnie liczy średnią i odchylenie standardowe ze wszystkich 5 etapów walidacji skrośnej.

Porównanie wyników

Pochodzenie danych	EEG						Peripheral					
	Arousal		Valence		Liking		Arousal		Valence		Liking	
	ACC	F1	ACC	F1	ACC	F1	ACC	F1	ACC	F1	ACC	F1
Artykuł	0.620	0.583	0.576	0.563	0.554	0.502	0.570	0.533	0.627	0.608	0.591	0.538
LSTM	0.476	0.528	0.524	0.619	0.602	0.722	0.500	0.555	0.548	0.642	0.610	0.726

W wierszu "Artykuł" podano wyniki uzyskane w oryginalnym artykule, a w wierszu "LSTM" podano wyniki uzyskane przez nas przy użyciu sieci rekurencyjnej LSTM.

W części przypadków nasz model daje lepsze wyniki niż model zaproponowany w artykule. Należy jednak zaznaczyć, że w walidacji skróśnej zaobserwowano odchylenie standardowe wyników dla poszczególnych części sięgające nawet $\pm 5\%$.

Wnioski

Wyniki są na tyle dobre, aby uznać, że sieć rekurencyjna jest dobrym modelem bazowym dla dalszych eksperymentów. Niepokojąca jest natomiast spora losowość wyników spowodowana najpewniej małą ilością danych uczących. Istotna może się okazać potrzeba zbadania stopnia losowości wyników, aby móc stwierdzić czy różnice pomiędzy wynikami różnych modeli nie są spowodowane przez losowość.