

Detecting Hand movements from EEG Signals*

Anirudhan J Rajagopalan, Michele Cerú†

Abstract. This project aims to detect and classify the grasp and lift hand movements of a subject from EEG (Electroencephalography) recordings. Successful identification and classification of the recordings will help in developing Brain Computer Interfaces (BCI) that can be used to restore the ability of patients to do day-to-day tasks. We are provided with 24 series of grasp and lift actions performed by twelve subjects. We run our models using the first 18 series as our training and the last 2 series as our test set. We use the level1 predictions of the Kaggle contest winners as our baseline and try to compare the performance of our model against the level1 predictions of baseline. Using our simple pipeline we are able to show that we can get accuracy of 0.71 Mean Area Under the Curve (MAUC).

1. Introduction. Electroencephalography (EEG) is used to record electrical activity of the brain. It is typically noninvasive, with the electrodes placed along the scalp. In the Grasp And Lift (GAL) experiment twelve subjects were asked to perform lifting series in which the object's weight (165, 330, or 660g), surface friction (sandpaper, suede, or silk surface), or both, were changed unpredictably between trials, thus enforcing changes in fingertip force coordination.[3] The hand movement of the subject was recorded by 3D sensors which were synchronized with the EEG cap thus providing us with the exact moment at which the GAL events happen. With respect to our goal of classifying the hand movements, we have to detect the following six events.

1. *HandStart*: Beginning of the movement.
2. *FirstDigitTouch*: Making contact with the object.
3. *BothStartLoadPhase*: Starting to load the object.
4. *LiftOff*: Holding the object up.
5. *Replace*: Replacing the object in its original position.
6. *BothReleased*: Releasing the fingers from the object.

The EEG signals recorded by 32 electrodes are recorded at a frequency of 500Hz. An added objective of this task is to make sure that we don't use any future data for doing predictions. This restriction is imposed to mimic the real life scenarios in which such an application can be used, wherein, we will not have access to any of the future data while making predictions in real life.[2]

This project aims to classify the hand movements of subjects by using EEG data. We compare the performance of SVM with Linear and Gaussian kernels by using VLAD and Bag of Features (BOF) feature representations. We try to reason the performance of our model with respect to the dataset by analysing the spatial relationship and variance of the features. We then discuss about models that we might use to tackle this hard problem.

2. Dataset Description.

*Kaggle Grasp and lift competition. <https://www.kaggle.com/c/grasp-and-lift-eeeg-detection>

†New York University (ajr619@nyu.edu; mc3784@nyu.edu). Questions, comments, or corrections to this document may be directed to this email address.

2.1. Sources and Format. The dataset is provided by Kaggle Inc. for their Grasp and Lift Detection challenge[2] sponsored by Way Consoritum[1]. The dataset consists of separate test and train zip files of size 153 MB and 915 MB respectively. The dataset expands to 447 MB and 3.1GB after extraction. Out of these datafiles, we cannot use the testset as they are being used for the competition submissions. So our effective usable dataset is only the training data which we partition into train and test set for training our models.

The dataset consists of two types of files:

1. data.csv — CSV of series labels and values from the 32 electrode signals.
2. event.csv — CSV of series labels and event values.

The sample dataset is included below for reference.

id	Fp1	Fp2	F7	F3	Fz	F4	F8	FC5	FC1	FC2	FC6	T7	C3	Cz	C4	T8	TP9	CP5	CP1	CP2	CP6	TP10	P7	P9	Pz	P4	P8	PO9	O1	Oz	O2	PO10
subj10_ser	-304	-156	-411	-640	-505	-603	-451	104	-344	-784	-387	-308	24	-725	-178	-335	-285	-936	-568	-378	-665	-107	-197	-587	-242	-472	-56	-338	-335	-518	-371	-177
subj10_ser	-300	-151	-376	-670	-509	-622	-422	60	-374	-815	-386	-326	11	-728	-189	-320	-307	-937	-569	-394	-651	-169	-176	-608	-224	-475	-78	-282	-289	-505	-424	-193
subj10_ser	-286	-227	-400	-724	-525	-611	-469	70	-378	-811	-390	-353	1	-738	-192	-343	-404	-907	-572	-389	-647	-179	-154	-605	-246	-488	-73	-304	-337	-539	-412	-192
subj10_ser	-281	-272	-447	-702	-540	-640	-492	71	-386	-807	-396	-354	5	-719	-167	-326	-373	-897	-549	-372	-637	-82	-146	-602	-226	-448	-25	-342	-333	-549	-400	-157
subj10_ser	-294	-194	-432	-693	-503	-609	-500	109	-385	-797	-372	-353	5	-719	-167	-326	-373	-897	-549	-372	-637	-82	-146	-602	-226	-448	-25	-342	-333	-549	-400	-157
subj10_ser	-297	-132	-406	-685	-526	-603	-521	144	-387	-810	-378	-360	-28	-741	-183	-333	-383	-950	-567	-377	-662	-104	-193	-609	-216	-448	-21	-378	-341	-582	-384	-152
subj10_ser	-286	-124	-372	-677	-531	-589	-442	162	-385	-804	-344	-325	0	-739	-160	-322	-403	-910	-562	-366	-623	-74	-200	-593	-205	-423	-19	-378	-354	-579	-404	-115
subj10_ser	-249	-131	-307	-630	-506	-588	-417	143	-366	-780	-324	-272	17	-691	-131	-310	-363	-865	-540	-318	-587	-31	-161	-563	-169	-379	20	-285	-264	-515	-335	-50
subj10_ser	-295	-171	-286	-667	-505	-616	-488	118	-377	-794	-329	-310	20	-693	-127	-266	-315	-859	-515	-297	-548	-46	-146	-531	-144	-336	39	-239	-249	-470	-312	-61

Figure 1. Sample Data format

id	HandStart	FirstDigitT	BothStart	LiftOff	Replace	BothReleased
subj10_ser	0	0	0	0	0	0
subj10_ser	0	0	0	0	0	0
subj10_ser	0	0	0	0	0	0
subj10_ser	0	0	0	0	0	0
subj10_ser	0	0	0	0	0	0
subj10_ser	0	0	0	0	0	0
subj10_ser	0	0	0	0	0	0
subj10_ser	0	0	0	0	0	0
subj10_ser	0	0	0	0	0	0

Figure 2. Sample Event format

The event dataset will have a value of 1 corresponding to the event (stimulus) that is being performed at the instant of sampling. There can be situations where there are multiple stimulus being triggered for the same EEG signal. The objective is to classify (or) label the input EEG signals to the corresponding stimulus.

The total number of samples across the complete dataset is 17985850.

2.2. Multilabeling or Multiclass classification. The dataset can have multiple stimulus at the same time as shown in figure 3. This is an important feature to consider as it forms the distinction between multiclass and multilabel classification problem.

```
subj1_series1_5254,0,1,1,1,0,0
subj1_series1_5255,0,1,1,1,0,0
subj1_series1_5256,0,1,1,1,0,0
subj1_series1_5257,0,1,1,1,0,0
subj1_series1_5258,0,1,1,1,0,0
subj1_series1_5259,0,1,1,1,0,0
subj1_series1_5260,0,0,1,1,0,0
```

Figure 3. *Multilabel or Multiclass classification problem.*

REFERENCES

- [1] WAY CONSORITUM, *Wearable interfaces for hand function recovery*. <http://www.wayproject.eu/>.
- [2] KAGGLE INC, *The home of data science*, June 2015. <https://www.kaggle.com/c/grasp-and-lift-eeg-detection>.
- [3] MATTHEW D LUCIW, EWA JAROCKA, AND BENONI B EDIN, *Multi-channel eeg recordings during 3,936 grasp and lift trials with varying weight and friction*, (2014). <http://dx.doi.org/10.1038/sdata.2014.47>.