Multi-task Learning for Commercial Brain Computer Interfaces

George Panagopoulos
gpanagopoulos@uh.edu , giorgospanagopoulos.github.io
Computational Physiology Lab ,University of Houston, Houston, Texas 77204

COMPUTATIONAL PHYSIOLOGY LAB

UNIVERSITY of HOUSTON
DEPARTMENT OF COMPUTER SCIENCE

Introduction

Problem

Improve subject generalization of passive economical Brain Computer Interfaces

Motivation:

Exploit the trade off between noisy individual recordings and increased number of subjects

Advantages:

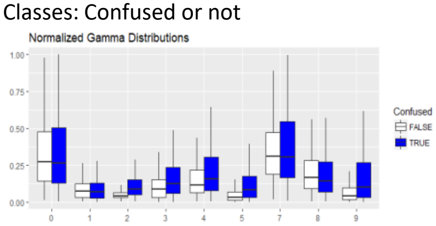
Utilize knowledge shared between subjects during training

Extract patterns present in all subjects instead of explicit individuals

Data

Berkeley Experiment [1]
- 30 subjects
- One 5-minute session each
- Two types of stimuli during session
  - Math, memorizing colors, think of items
  - Listen to music, watch video ads, relax
- Cognitive state changes in the same session
- Classes: Mental activity or relaxation

Carnegie Mellon Experiment [2]
- 9 subjects
- Ten 2-minute sessions with MOOC videos
- Self-classified levels of confusion for each session
- Cognitive state changes between sessions
- Classes: Confused or not



Example of subject variability

Methods

Discriminative MTL[1]

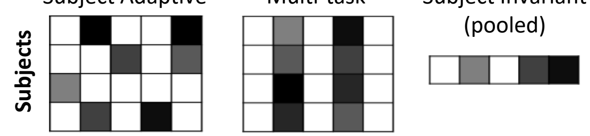
Logistic regression with group sparsity constraint on the K coefficients of all subjects T

$$min\frac{1}{N_t}\sum_{t=1}^{T} J(X_t, W_t, Y_t) + \lambda \sum_{k=1}^{K} |W_k|_2$$

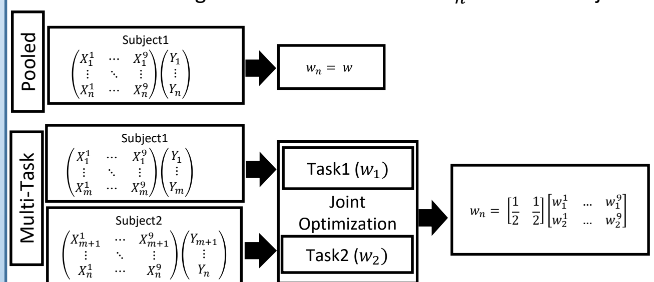
Rows of W are the feature coefficients w_t of each subject t

Each approach yields a different W:

Feature Coefficients



The column average of multi-task W is the w_n of a new subject n



Generative MTL[2]

Bayesian estimation of coefficients' prior distribution

$$p(W;X,Y)\sim\prod_{t=1}^{T} p(y_t, X_t; w_t)p(w_t)$$

$$p(w_t)\sim N(\mu,\Sigma)\ ,\forall\ t \in T$$

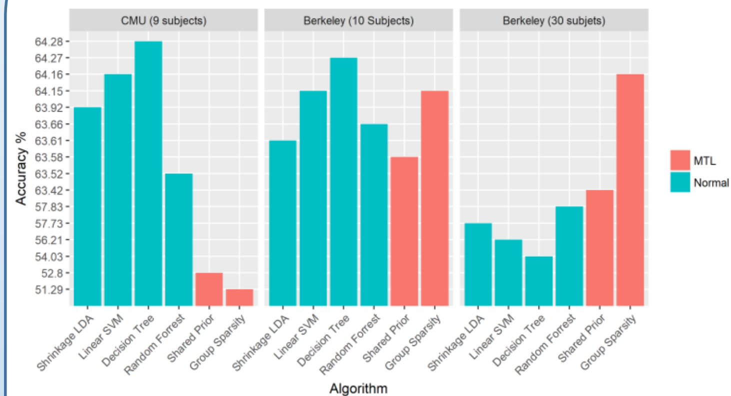
$$\min_{w_t,\mu,\Sigma}\frac{1}{\sigma^2}\sum_{t=1}^{T}\left||X_t w_t - Y_t|\right|^2 + \frac{1}{2}\sum_{t=1}^{T}(w_t-\mu)^T\Sigma^{-1}(w_t-\mu) + \frac{T}{2}logdet(\Sigma)$$

The sign of the regressed value determines the binary class

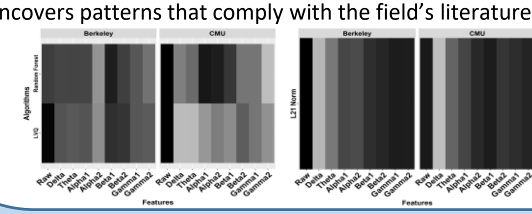
For a new subject a coefficient vector w_n is sampled from $N(\mu,\Sigma)$

Results

MTL methods perform better as the number of subjects increases



MTL uncovers patterns that comply with the field's literature



Conclusion

Multi-task algorithms are more robust than conventional pooled approaches to the subject generalization problem.

- Improved or steady accuracy with more subjects
- Consistent among datasets feature selection

References

[1] Argyriou, Andreas, Theodoros Evgeniou, and Massimiliano Pontil. "Convex multi-task feature learning." Machine Learning 73.3 (2008): 243-272.

[2] Alamgir, Morteza, Moritz Grosse-Wentrup, and Yasemin Altun. "Multitask learning for brain-computer interfaces." International Conference on Artificial Intelligence and Statistics. 2010.

Code

https://github.com/GiorgosPanagopoulos/Multi-task-Learning-for-Commercial-Brain-Computer-Interfaces

[1] https://www.kaggle.com/berkeley-biosense/synchronized-brainwave-dataset [2] https://www.kaggle.com/wanghaohan/eeg-brain-wave-for-confusion