Inter-session Transfer Learning using Riemannian Geometry for Mental Workload Estimation

Divyesh Narayanan[1]

[1]Department of Electronics and Instrumentation Engineering, Birla Institute of Technology and Science, Pilani

EEG signals are highly non-stationary and can drift in terms of amplitude and other features over time, even within a day for a particular user. So, there is a need for algorithms capable of inter-session transfer learning such that the calibration is minimized or eliminated for the next session. This abstract outlines the method used to do a 3-class workload estimation in an inter-session, intra-subject manner on the dataset released as a part of the Passive BCI Hackathon conducted with the Neuroergonomics Conference 2021[1].

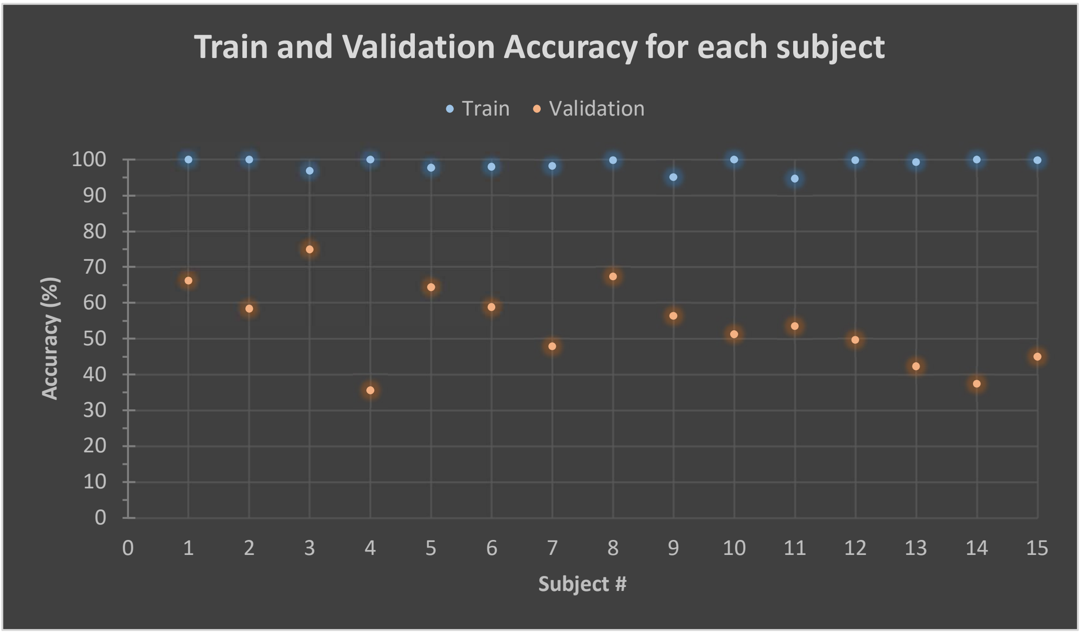
The data preprocessed for noise rejection is provided as 2-second epochs at 250Hz. Covariance matrices are Symmetric Positive Definite matrices. Hence, the correct manipulation, which best uses the information in these matrices, is based on a branch of differential geometry, i.e., Riemannian geometry [2]. For each point C (the covariance matrix) of the Riemannian manifold, the tangent space is Euclidean and locally homomorphic to the manifold, and Euclidean distance computations can reasonably approximate the Riemannian distance computations in the manifold in the tangent space. The Riemannian distance (d) between 2 covariance matrices, A and B, is calculated as the root of the sum of the log of the joint eigenvalues of A and B squared and the Riemannian geometric mean of covariance matrices (also known as Fréchet mean) is calculated as the matrix minimizing the sum of the square Riemannian distances [3,4].

The ensemble model built is a soft voting classifier with equal weights that adds the prediction probabilities of three Riemannian geometry-based models to make predictions based on the class with maximum prediction probability. The preprocessed data is used to estimate the covariance across channels and is given as the input feature to the models. The first model uses a Minimum Distance to Mean classifier. The centroid (mean covariance) is estimated for each class according to the chosen metric, the Riemannian distance. For each new point, the distance of the covariance matrix from each of the centroids is calculated using the Riemannian distance metric, and the class is chosen based on the nearest centroid. For the second model, the covariance matrix is projected to the tangent space, and a support vector classifier is used. This model makes use of an adaptive kernel to account for inter-session variability [5]. It uses the prediction data in an unsupervised manner to update the reference point (the Arithmetic/Geometric mean) for the tangent space to which the new data are projected. The classifier is used with unchanged parameters across sessions. For the third model, the same adaptive projection method to the tangent space and a tuned gradient boosting classifier is used. For reproducibility, the codebase is available at https://github.com/Div12345/WorkloadEstimation

Training on the second session gives better generalization and prediction on the first session data based on some preliminary runs, and so this was chosen as the measure of validating the performance of the algorithms. The performance of the model on the different subjects is shown in Fig 1. Apart from Subjects 4 and 14, the accuracy was above 40% for all subjects on the validation data. 4 subjects had more than 60% validation accuracy, including one at 74.94%. The overall average accuracy across subjects on the validation data is 51.25%. The results show significant and well above chance (33%) inter-session classification for all the subjects.

Passive brain signal decoding is especially hard considering the lack of event markers and labeled datasets. The dataset used here provides a unique opportunity to understand and solve this problem. This work shows that the ensembling of techniques using the fundamental properties of covariance matrices through Riemannian geometry combined with unsupervised transfer learning using an adaptive kernel for tangent space projection shows a promising result for inter-session generalization. Though the transfer learning technique used here isn't particularly original, it is necessary to attempt pre-established strategies on new datasets before developing new techniques. This work should serve as a baseline for more specialised techniques developed in the future.

Fig 1. Performance of the model. Shows the accuracy (in %) for each subject in training (session 2) and testing (session 1)



References:

[1] Hinss, Marcel F., Darmet, Ludovic, Somon, Bertille, Jahanpour, Emilie, Lotte, Fabien, Ladouce, Simon, & Roy, Raphaëlle N. (2021). An EEG dataset for cross-session mental workload estimation: Passive BCI competition of the Neuroergonomics Conference 2021 (Version 2) [Data set]. Neuroergonomics Conference, Munich, Germany. Zenodo. https://doi.org/10.5281/zenodo.5055046

[2] Berger, Marcel. (2003). A Panoramic View of Riemannian Geometry. 10.1007/978-3-642-18245-7_13.

[3] Moakher, Maher. (2005). A Differential Geometric Approach to the Geometric Mean of Symmetric Positive-Definite Matrices. SIAM J. Matrix Analysis Applications. 26. 735-747. 10.1137/S0895479803436937.

[4] Pennec, X., Fillard, P. & Ayache, N. A Riemannian Framework for Tensor Computing. *Int J Comput Vision* **66,** 41–66 (2006). https://doi.org/10.1007/s11263-005-3222-z

[5] Barachant, A., S. Bonnet, M. Congedo and C. Jutten. "Classification of covariance matrices using a Riemannian-based kernel for BCI applications." *Neurocomputing* 112 (2013): 172-178.