**Project Specification - Degree Project in Computer Science, DD142x**

## Generalisation of brain signal classification machine learning models from single subjects to multiple subjects

March 12, 2018, Stockholm

Axel Karlsson and Victor Wiklund

Supervisor: Pawel Herman

## Introduction

In the last few years the potential of machine learning has been realised in many fields such as speech recognition [1], image recognition [2] and even brain signal pattern recognition [3]. In order for a machine learning implementation to be successful a large amount of data is required. In the case of brain signal pattern recognition a common method for obtaining brain signal data is using electroencephalography(EEG). EEG is a typically noninvasive method where electrodes are placed on the head of the subject. The electrodes register the electrical signals sent by the brain and send these to a computer which interprets the signals. Creating a machine learning model with EEG data is in turn a possibility for the creation of a brain-computer interface(BCI) [4]. A BCI is a tool which enables direct communication between a computer and a brain with several interesting potential applications such as new ways of controlling prosthetic limbs or interfacing with machines.

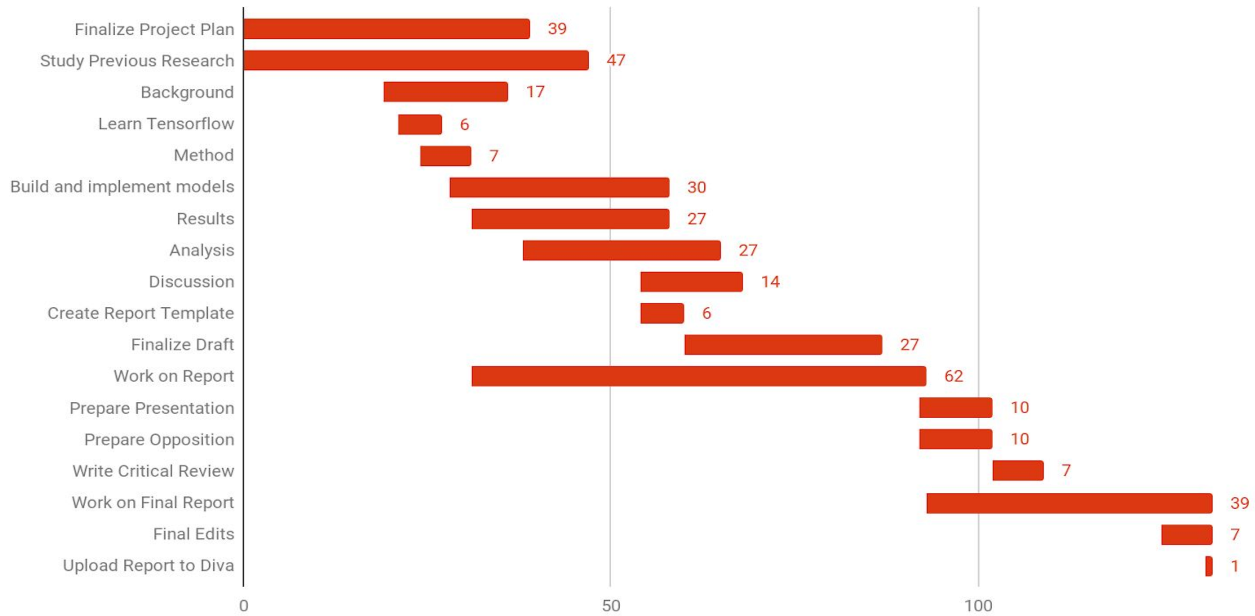
## Problem Statement

One challenging aspect of using machine learning as a tool for interpretation is the difficulty of tuning the model to maximize performance - especially so in EEG-classification where inter-subject variability is abundant [5]. Understanding how to modify a model to account for new subject data without losing performance is thus of central importance. To that end we pose the question; How does inter-subject generalization compare to subject cross-session generalization?

## Approach

The datasets will be split into datasets for individual subjects and smaller groups of subjects including all the data for the relevant subjects across all sessions. For each dataset we will build and optimize three machine learning models. Two of these models will be Artificial Neural Networks (ANN) with one being a convolutional neural network(CNN) and the other being a Recurrent neural network(RNN). The last model will be a Support vector machine(SVM). We chose CNNs because hypothetically they should have good results for cross-subject generalization due to their powerful generalisation and invariance handling capabilities[6, 7]. We chose RNNs because like CNNs RNNs are commonly used deep neural networks [1, 8] but which also respect the time aspect of the data. We believe that the comparison between these different types of deep neural networks will give valuable insight. Lastly we also chose the SVM method in order to gain some insight into generalisation for a machine learning approach that is popular in BCI research [9, 10]. To build the network and manage the hyperparameters we will use Tensorflow [REF]. We will rely on Tensorboard to visualize and compare the networks performance on different data. The data we will work with is preprocessed so we will not invest efforts into feature extraction and other preprocessing routines. To optimize the neural networks with respect to the hyperparameters we will employ Bayesian Optimization and Gaussian Processes because these optimization methods have successfully been used for hyperparameter tuning of deep neural network to find better hyperparameters faster than other global optimization algorithms [12]. To optimize the SVM with respect to the hyperparameters we will apply estimation distribution algorithms since they are generally capable of reaching better results than random search [13]. In order to conduct null hypothesis and ANOVA testing when comparing the results from the models the Numpy and Scipy python libraries will be utilized. We will compare the performance of the different models and also try to tune the models that use the inter-subject data so that their accuracy when classifying multiple subjects is maximized.

## Timeplan




## References

[1] A. Graves, A.-R. Mohamed, and G. Hinton, “Speech recognition with deep recurrent neural networks,” in *2013 IEEE International Conference on Acoustics, Speech and Signal Processing*, 2013.

[2] A. Krizhevsky, I. Sutskever, and G. E. Hinton, “ImageNet classification with deep convolutional neural networks,” *Commun. ACM*, vol. 60, no. 6, pp. 84–90, 2017.

[3] J. Zhang, C. Yan, and X. Gong, “Deep convolutional neural network for decoding motor imagery based brain computer interface,” in *2017 IEEE International Conference on Signal Processing, Communications and Computing (ICSPCC)*, 2017.

[4] S. N. Abdulkader, A. Atia, and M.-S. M. Mostafa, “Brain computer interfacing: Applications and challenges,” *Egyptian Informatics Journal*, vol. 16, no. 2, pp. 213–230, 2015.

[5] S. N. G. Bolagh, M. B. Shamsollahi, C. Jutten, and M. Congedo, “Unsupervised Cross-Subject BCI Learning and Classification using Riemannian Geometry,” in *24th European Symposium on Artificial Neural Networks, Computational Intelligence and Machine Learning (ESANN 2016)*, 2016.

[6] R. T. Schirrmeister *et al.*, “Deep learning with convolutional neural networks for EEG decoding and visualization,” *Hum. Brain Mapp.*, vol. 38, no. 11, pp. 5391–5420, Nov. 2017.

[7] S. Lawrence, C. L. Giles, Ah Chung Tsoi and A. D. Back, "Face recognition: a convolutional neural-network approach," in *IEEE Transactions on Neural Networks*, vol. 8, no. 1, pp. 98-113, Jan 1997.

[8] A. Petrosian, D. Prokhorov, R. Homan, R. Dasheiff, and D. Wunsch, “Recurrent neural network based prediction of epileptic seizures in intra- and extracranial EEG,” *Neurocomputing*, vol. 30, no. 1–4, pp. 201–218, 2000.

[9] I. Güler and E. D. Ubeyli, “Multiclass support vector machines for EEG-signals classification,” *IEEE Trans. Inf. Technol. Biomed.*, vol. 11, no. 2, pp. 117–126, Mar. 2007.

[10] A. Subasi and M. Ismail Gursoy, “EEG signal classification using PCA, ICA, LDA and support vector machines,” *Expert Syst. Appl.*, vol. 37, no. 12, pp. 8659–8666, 2010.

[11] M. Abadi, A. Agarwal, P. Barham, E. Brevdo, Z. Chen, C. Citro, G. S. Corrado, A. Davis, J. Dean, M. Devin, S. Ghe- mawat, I. Goodfellow, A. Harp, G. Irving, M. Isard, Y. Jia, R. Jozefowicz, L. Kaiser, M. Kudlur, J. Levenberg, D. Mane ́, R. Monga, S. Moore, D. Murray, C. Olah, M. Schuster, J. Shlens, B. Steiner, I. Sutskever, K. Talwar, P. Tucker, V. Vanhoucke, V. Vasudevan, F. Vie ́gas, O. Vinyals, P. War- den, M. Wattenberg, M. Wicke, Y. Yu, and X. Zheng. Tensor- Flow: Large-scale machine learning on heterogeneous sys- tems, 2015. Software available from tensorflow.org.

[12] Wenqiang Guo, W. Guo, X. Gao, and Q. Xiao, “Bayesian Optimization Algorithm for learning structure of dynamic bayesian networks from incomplete data,” in *2008 Chinese Control and Decision Conference*, 2008.

[13] L. C. Padierna, M. Carpio, A. Rojas, H. Puga, R. Baltazar, and H. Fraire, “Hyper-Parameter Tuning for Support Vector Machines by Estimation of Distribution Algorithms,” in *Studies in Computational Intelligence*, 2016, pp. 787–800.