Machine Learning Methods for Computational Psychology

Sarah M. Brown
Northeastern University, Boston, MA
Charles Stark Draper Laboratory, Cambridge, MA

Abstract

Advances in sensing and imaging have provided psychology researchers new tools to understand how the brain creates the mind and simultaneously revealed the need for a new paradigm of mind-brain correspondence– a set of basic theoretical tenets and an overhauled methodology. I develop machine learning methods to overcome three initial technical barriers to application of the new paradigm. I assess candidate solutions to these problems using two test datasets representing different areas of psychology: the first aiming to build more objective Post-Traumatic Stress Disorder (PTSD) diagnostic tools using virtual reality and peripheral physiology, the second aiming to verify theoretical tenets of the new paradigm in a study of basic affect using functional Magnetic Resonance Imaging (fMRI). Specifically I address three technical challenges: assessing performance in small, real datasets through stability; learning from labels of varying quality; and probabilistic representations of dynamical systems.

Introduction

Improvements in neuroimaging technology enable psychologists to consider how the brain creates the mind with a degree of complexity that has been inaccessible to date. Increasingly, however, existing theory is unable to explain the observed phenomena, revealing the need for a new paradigm: core theoretical tenets and methodology. In a joint research effort we have developed a candidate paradigm for psychology research, the Unified Framework for Brain-Mind Modeling (UFBMM) (Barrett et al. 2015). In my thesis, I develop machine learning solutions to enable preliminary analysis of data in light of the new paradigm. First we consider the task of re-analyzing old, likely undersized data sets for preliminary result and to guide experimental design. Second, we take such an intermediate step in a Post Traumatic Stress Disorder diagnosis setting using traditional features but relaxing constraints and interpreting the result as a proof of preliminary concept. Finally, we consider time series analysis in light of some of the core theoretical tenets of the new paradigm.

We approach these problems through two specific experimental data sets. The first is an exploratory experiment toward a PTSD Diagnosis system that scores patients while they view Virtual Reality videos using peripheral (distant from the brain) physiological measures. The second data set considers basic affect, but innovates from standard fMRI data sets by presenting subjects with 900 images instead of the typical 30 to enable the study of dynamics and interactions. The solutions provided by algorithms must be understood by clinical and psychological researchers to enable further hypothesizing and support decision making.

Stability as a Performance Measure

Re-analysis of existing datasets in light of new theoretical tenets is a convenient method for validating the new paradigm and generating preliminary results to facilitate design of new experiments. These datasets may be statistically small relative to the class of analysis techniques suggested by the new paradigm, so differentiating between a weak signal and an overfit model is an important capability for promotion of the new paradigm. Intuitively motivated, stability has served as a useful heuristic for similar settings in unsupervised learning (Von Luxburg, Ben-david, and Luxburg 2005) and feature selection (Kalousis, Prados, and Hilario 2005). We relate these empirical applications as a performance measure to theoretical results for stability as a guarantee for good generalization (Bousquet and Elisseeff 2002) to position stability as a framework for deriving application-appropriate performance measures. First, I design a general framework for defining and comparing machine learning applications of stability and show how to recover definitions from the literature. This framework reveals necessary analytical gaps to extend and relate the theoretical and practical uses, thus enabling theoretical analysis of the empirical applications. I am working to fill these theoretical gaps and anticipate results by February. I will apply stability-based performance metrics derived from this template to each of the subsequent analyses.

Learning from a catch-all label

Virtual reality has been used for PTSD treatment (Rizzo et al. 2010) and a number of physiological measures have been shown to vary with PTSD diagnosis (Orr, Metzger, and Pitman 2002; Webb et al. 2013). Combined, these technologies present an opportunity for closed-loop treatment monitoring: updated diagnosis during the treatment protocol. A necessary prerequisite to such a diagnostic system is a physiologically computed PTSD score that agrees with the clinical gold standard. The gold standard diagnostic tool, Clinician Administered PTSD Scale (CAPS) is a structured clinical interview designed to rank people above or near a threshold, and assigns a zero to patients with no symptoms (Blake et al. 1995; American Psychiatric Association 2013). Therefore, the CAPS can be interpreted as a severity *score* (1-140) or a healthy *label* (0). This mixed composition poses a problem for standard regression models, so we develop a novel learning formulation, Sparse Combined Regression-Classification (SCRC) (Brown et al. 2015), to learn a function from this ambiguous training data that computes a severity score from standard psycho-physiological features. A scoring function learned with SCRC outperforms that of a naive computational method on metrics designed to assess diagnostic validity, parsimony, and generalizability.

Time Series Analysis

The UFBMM posits a state space model for the brain where physiological and behavioral measurements are partial observations of the brain state. The two groups of measurements share some information, but are not equivalent or even noisy representations of one another. We therefore formulate analysis of psychological experimental data as an unsupervised, multi-observation task. I propose Bayesian nonparametric models for structure discovery (Duvenaud et al. 2013; Fox and Dunson 2012; Saria, Koller, and Penn 2010) and modeling latent component structure (Griffiths and Ghahramani 2011) to interpret data. To capture temporal structure we leverage the expressivity and flexibility of Gaussian Processes. For the PTSD dataset, I will develop a multichannel model that discovers latent components that each express in a subset of the measurement channels. For the fMRI dataset, in a collaborative effort, we will add spatial and network structure to the physiological observation model. The resultant models will identify latent components that express both physiologically and behaviorally, without enforcing a causal relationship from one measurement to the other.

Conclusion

My thesis provides machine learning methods catered to answering a new class of research questions in psychology. I develop methods that are context sensitive for extracting insight from psychology experimental data, by adopting a model based approach with simple recommendations in order to promote interpretable solutions. Motivated by practical challenges associated with interpreting data in a transient mode of science, I provide a framework for design of context-appropriate performance measures that can be reused in a broad variety of applications. I formulated an application appropriate task and leveraged features familiar to clinicians to produce an interpretable proof of concept score in physiological PTSD diagnosis. After a collaborative effort to design an abstract mathematical model for a new theory of brain-mind mapping, I have developed computational techniques that allow for re-analysis of experiments designed in the current tradition for interpretation in the context of a new paradigm.

Acknowledgments

This work was supported by a National Science Foundation Graduate Research Fellowship (NSF DGE-0946746) and Charles Stark Draper Laboratory.

References

American Psychiatric Association. 2013. The Diagnostic and Statistical Manual of Mental Disorders: DSM 5. Technical report, American Psychiatric Association.

Barrett, L. F.; Brooks, D.; Brown, S. M.; Clark-Polner, E.; Coll-Font, J.; Dy, J. G.; Erdogmus, D.; Erem, B.; Satpute, A. B.; and Wilson-Mendenhall, C. D. 2015. Inferring the Mind by Modelling the Brain : Beyond Faculty Psychology. *Manuscript in Progress*.

Blake, D. D.; Weathers, F. W.; Nagy, L. M.; Kaloupek, D. G.; Gusman, F. D.; Charney, D. S.; and Keane, T. M. 1995. The development of a clinician-administered PTSD scale. *Journal of traumatic stress* 8(1):75–90.

Bousquet, O., and Elisseeff, A. 2002. Stability and Generalization. *The Journal of Machine Learning Research* 2:499–526.

Brown, S. M.; Webb, A.; Mangoubi, R. S.; and Dy, J. G. 2015. A Sparse Combined Regression-Classification Formulation for Learning a Physiological Alternative to Clinical Post-Traumatic Stress Disorder Scores. In *Association for the Advancement of Artificial Intelleigence 2015*.

Duvenaud, D.; Lloyd, J.; Grosse, R.; Tenenbaum, J.; and Ghahramani, Z. 2013. Structure discovery in nonparametric regression through compositional kernel search. In *Proceedings of the International Conference on Machine Learning (ICML)*, volume 30, 1166–1174.

Fox, E. B., and Dunson, D. B. 2012. Multiresolution Gaussian Processes. In Pereira, F.; Burges, C. J. C.; Bottou, L.; and Weinberger, K. Q., eds., *Advances in Neural Information Processing Systems 25*, 737–745. Curran Associates, Inc.

Griffiths, T. L., and Ghahramani, Z. 2011. The Indian Buffet Process: An Introduction and Review. *Journal ofMachine Learning Research* 12:1185–1224.

Kalousis, A.; Prados, J.; and Hilario, M. 2005. Stability of feature selection algorithms. In *Data Mining, Fifth IEEE International Conference on*, 8—-pp. IEEE.

Orr, S. P.; Metzger, L. J.; and Pitman, R. K. 2002. Psychophysiology of post-traumatic stress disorder. *The Psychiatric clinics of North America* 25(2):271–93.

Rizzo, A.; Difede, J.; Rothbaum, B. O.; Reger, G.; Spitalnick, J.; Cukor, J.; Mclay, R.; and Others. 2010. Development and early evaluation of the Virtual Iraq/Afghanistan exposure therapy system for combat-related PTSD. *Annals of the New York Academy of Sciences* 1208(1):114–125.

Saria, S.; Koller, D.; and Penn, A. 2010. Discovering shared and individual latent structure in multiple time series. *arXiv preprint arXiv:1008.2028* stat.ML(d):1–9.

Von Luxburg, U.; Ben-david, S.; and Luxburg, U. V. 2005. Towards a statistical theory of clustering. In *Pascal workshop on statistics and optimization of clustering*, 20–26.

Webb, A. K.; Vincent, A. L.; Jin, A.; and Pollack, M. H. 2013. Wearable sensors can assist in PTSD diagnosis. *2013 IEEE International Conference on Body Sensor Networks* 1–6.