Research Statement

of Shreya Kapoor (PhD applicant for Fall—2020)

My primary research interest lies at the intersection of Neuroscience and Machine learning with a focus on Neural Data analysis. I want to orient my prospective PhD towards Theoretical Neuroscience with an emphasis on Machine Learning. Through the course of the doctorate I want to develop models for understanding neural responses in diseased states. Task-based fMRI is used to detect presurgical motor, language, and memory mapping in subjects. It's spatiotemporal nature gives a measure of activity of the brain while performing an action but it remains unclear how the obtained data features encode behavioural information. Further, the relationships between input data and the behaviour exhibited by subjects while performing a task captured by Machine learning are usually mathematical functions which do not reflect biological mechanisms perfectly. Hence, I am particularly interested in methods of improving the interpretability of machine learning model predictions on task-based fMRI data for medical diagnosis.
Machine learning is often applied to major task based fMRI studies in order to classify disease states of subjects. Classification is based on inter-subject differences in haemodynamic response and these differences are considered to be correlated to the cause of a disease. To meet the criterion of significant confidence level for clinical decision-making, it is important to enhance the interpretability of machine learning model predictions. fMRI gets temporal activity from thousands of voxels and most discriminative approaches rely on spatial information or regional correlations but do not highlight the mechanism behind any abnormality[3]. Traditional methods such as univariate analysis are not suited for such applications since purely mathematical measures such as correlations do not provide any Neurobiological insight. The problems that I would like to bring into consideration while inferring from neural data are: How can we encode massive amounts of data in a less complex representation, and how well does this representation form a feature space for machine learning models to detect interpretable patterns?
The first problem needs to be addressed since fMRI data is often prone to a low signal-to-noise ratio and a high degree of disuniformity between subjects which make obtaining a comparatively simple representation with rich information content a difficult task. Formulating hypothesis from spatial distribution could lead to inaccuracies. Various experiments have shown that the spatial distribution of activity is not the only factor responsible for eliciting neural responses[4]. It has been observed that hidden physiological quantities, such as synaptic connection weights, might be more illustrative to highlight the same activity. Further, parameters which cannot be measured directly are important for information processing in a biological system; Panzeri et al.[6] modelled sensory perception using neural population features that are low-dimensional projections of large population activity. Understanding the abstraction underlying a mathematical representation of neural activity is useful since the results from discriminative models are artefacts of it.

The second problem follows from the first one: How well would a data representation reduced through Neurobiological insights fare in a purely mathematical Machine Learning model? For example, if one takes a lower-dimensional projection from a region of interest including multiple voxels but without any properly defined mapping of biological correspondence the explainability of the results suffer. To be able to design a feature space Neurobiology is necessarily needed to be taken into account, such is the case when machine Learning models have been used to predict behaviour of subjects and have demonstrated the ability to 'read' information from fMRI scans but replicability of these results is debatable [1]. For instance, Mitchell et al. [5] proposed techniques to model responses to a variety of stimuli, such as the semantic category of words presented to test subjects. Such models look for informative features that illustrate how the brain encodes states and stimuli.

As a future line of research, I would like to investigate methods of embedding map voxel-based data to deterministic feature space in an unsupervised manner and represent conditional dependencies of behavior on population coding. Previous research has shown such a mapping can be achieved using Generative Embeddings of fMRI data[2]. Also, conditional probabilistic models can represent the dependencies of behavior/labels on neural responses in reduced feature space [7]. Since these two methodologies could possibly answer the two research questions I want to investigate respectively, it would be interesting to combine them in order to design graphical networks which offer an insight into mechanisms of diseases. Designing such a pipeline could guide knowledge discovery about differences in behavior with respect to disease state.

My motivation to pursue research at the interface of machine learning and neuroscience has matured during my master studies. I particularly enjoyed taking a course on Visual Computing in the Life Sciences given by Prof. Thomas Schultz. It focused on graphical representations of high dimensional data and pipelines for fMRI image processing including image segmentation, registration and statistical analysis. I learnt that for such tasks it's important to do backward modelling i.e. determining the goal of the predictions and then designing the knowledge discovery approach with feedback loops. During this course, I implemented a variety of tasks for neural image processing, throughout the pipeline it was important to be able to explain what each transformation meant visually and mathematically.

In my view, my area of interest would align well with the interdisciplinary research at the Gatsby unit since there a strong focus on formulating a mathematical framework on neuroscience problems. For the neuroscientific component of my research I find the conceptual framework proposed by Prof. Peter E Latham and his colleagues as a good starting point/inspiration about the intersection between sensory coding and information for my prospective research, I am intrigued by how they highlight neural coding features as an intersection between sensory and choice information[6]. Prof Arthor Gretton's work on usage of kernel methods to drive inference in graphical models, by learning from the data in an unsupervised manner could help address my second research question about giving input to machine learning models from a reduced feature space.

One of the main reasons I find the Gatsby unit my first choice for pursuing doctoral research is strength of its academic collaborations all over the world. I am particularly excited about the Unit's proximity to the Wellcome Centre for Human Neuroimaging since Machine learning insights could help the Centre accomplish its goal to understand how behavior arises from activity.

References

[1] Rotem Botvinik-Nezer, Roni Iwanir, Felix Holzmeister, Jürgen Huber, Magnus Johannesson, Michael Kirchler, Anna Dreber, Colin F Camerer, Russell A Poldrack, and Tom Schonberg. fMRI data of mixed gambles from the Neuroimaging Analysis Replication and Prediction Study. *Scientific Data*, 6(1): 106, 2019. ISSN 2052-4463. doi: 10.1038/s41597-019-0113-7. URL `https://doi.org/10.1038/s41597-019-0113-7`.

[2] Kay H Brodersen, Thomas M Schofield, Alexander P Leff, Cheng Soon Ong, Ekaterina I Lomakina, Joachim M Buhmann, and Klaas E Stephan. Generative Embedding for Model-Based Classification of fMRI Data. *PLOS Computational Biology*, 7(6):e1002079, jun 2011. URL `https://doi.org/10.1371/journal.pcbi.1002079`.

[3] John-Dylan Haynes and Geraint Rees. Predicting the orientation of invisible stimuli from activity in human primary visual cortex. *Nature Neuroscience*, 8 (5):686–691, 2005. ISSN 1546-1726. doi: 10.1038/nn1445. URL `https://doi.org/10.1038/nn1445`.

[4] Yukiyasu Kamitani and Frank Tong. Decoding the visual and subjective contents of the human brain. *Nature Neuroscience*, 8(5):679–685, 2005. ISSN 1546-1726. doi: 10.1038/nn1444. URL `https://doi.org/10.1038/nn1444`.

[5] Tom M Mitchell, Rebecca Hutchinson, Marcel A Just, Radu S Niculescu, Francisco Pereira, and Xuerui Wang. Classifying instantaneous cognitive states from FMRI data. *AMIA ... Annual Symposium proceedings. AMIA Symposium*, 2003:465–469, 2003. ISSN 1942-597X. URL `https://www.ncbi.nlm.nih.gov/pubmed/14728216https://www.ncbi.nlm.nih.gov/pmc/articles/PMC1479944/`.

[6] Stefano Panzeri, Christopher D Harvey, Eugenio Piasini, Peter E Latham, and Tommaso Fellin. Cracking the Neural Code for Sensory Perception by Combining Statistics, Intervention, and Behavior. *Neuron*, 93(3):491–507, feb 2017. ISSN 0896-6273. doi: 10.1016/j.neuron.2016.12.036. URL `https://doi.org/10.1016/j.neuron.2016.12.036`.

[7] T P Speed and H T Kiiveri. Gaussian Markov Distributions over Finite Graphs. *Ann. Statist.*, 14(1):138–150, 1986. ISSN 0090-5364. doi: 10.1214/aos/1176349846. URL `https://projecteuclid.org:443/euclid.aos/1176349846`.