Detection of Cognitive States from fMRI data using Machine Learning Techniques

Vishwajeet Singh, K.P. Miyapuram, Raju S. Bapi*
University of Hyderabad
Computational Intelligence Lab, Department of Computer and Information Sciences
Gachibowli, Hyderabad, India 500046
bapics@uohyd.ernet.in

Abstract

Over the past decade functional Magnetic Resonance Imaging (fMRI) has emerged as a powerful technique to locate activity of human brain while engaged in a particular task or cognitive state. We consider the inverse problem of detecting the cognitive state of a human subject based on the fMRI data. We have explored classification techniques such as Gaussian Naive Bayes, k-Nearest Neighbour and Support Vector Machines. In order to reduce the very high dimensional fMRI data, we have used three feature selection strategies. Discriminating features and activity based features were used to select features for the problem of identifying the instantaneous cognitive state given a single fMRI scan and correlation based features were used when fMRI data from a single time interval was given. A case study of visuo-motor sequence learning is presented. The set of cognitive states we are interested in detecting are whether the subject has learnt a sequence, and if the subject is paying attention only towards the position or towards both the color and position of the visual stimuli. We have successfully used correlation based features to detect position-color related cognitive states with 80% accuracy and the cognitive states related to learning with 62.5% accuracy.

1 Introduction

Functional Magnetic Resonance Imaging (fMRI) is a powerful imaging tool, which can be used to perform brain activation studies my measuring Blood Oxygen Level Dependent (BOLD) signal [Ogawa et al., 1990]. Human brain carries out various functions at a time. The operations within the brain like, "the process of thinking" and "the process of remembering", that affect our mental contents are called as 'cognitive processes' and the person is said to be in the particular 'cognitive state'. Our aim is to identify the cognitive state of a human subject that is persistent within a certain interval of time, given the fMRI activity within that interval. Our motivation in the detection of cognitive states is that this could lead to our understanding of the hidden cognitive states a human passes through while performing a particular task. This could have potential applications in Lie Detectors, Cognitive control of artifacts, Cognitive control of mental states (Neurofeedback), etc.

An fMRI scanner measures the value of the fMRI signal at all the points in a three dimensional grid, or image every few seconds (6 sec in our data). The voxels (volume pixels) in a typical fMRI study have a volume of a few tens of cubic millimeters and the 3D brain image contains 1,84,707 voxels that constitute the whole brain, thereby resulting in very high dimensional data. Since the signals measure tiny fluctuations in the magnetic field, known as the Blood Oxygen Level Dependent (BOLD) response, the signal-to-noise ratio (SNR) for fMRI data is very low. Thus the other characteristic of fMRI data is that they are very noisy. This problem domain is quite interesting from the Machine Learning perspective, as it provides a case study of classifier learning from extremely high dimensional and noisy data.

We borrow Mitchell's [Mitchell, 1997] definition to state more formally the problem of detection of cognitive states from fMRI data.

Definition: A computer program is said to have learned from experience E with respect to some class of tasks T and performance measure P, if its performance at tasks in T, as measured by P, improves with experience E. In our case,
T: classification of cognitive state of human subject
E: labelled training data mapping fMRI scans to the instantaneous cognitive state
P: performance accuracy of the algorithm over the unseen fMRI data.

The case study in this paper is a Visuomotor sequence learning experiment [Miyapuram, 2004]. Human Subjects perform a motor task (pressing a key on a keyboard) in response to visual stimuli (having position and colour attributes). The problem that we have attempted in this study is to classify fMRI data into a discrete set of cognitive states that are "subject has learnt a sequence", "subject did not learn the sequence", "the subject is paying attention only towards the position of the visual stimuli" and "the subject is paying attention towards the color and the position of the visual stimuli". The various machine learning algorithms we have applied on this fMRI data are Gaussian Naive Bayes Classifier (GNB), k-Nearest Neighbor (kNN) and Support Vector Machines (SVM). To deal with the high dimensionality of the data we have used three feature selection strategies namely, Discriminating features, Activity based features and Correlation based features.

2 Materials and Methods

2.1 Machine Learning Classifier Techniques

A classifier is a function which takes input a set of examples and outputs a class label to each input example, where class label belongs to a discrete set of possible categories. In our case the classifier is a function of the form:
f : fMRI-sequence(t_1, t_2) → CognitiveState
where fMRI-sequence(t_1, t_2) is the sequence of fMRI images collected during the contiguous time interval [t_1, t_2], and where CognitiveState is the set of cognitive states to be discriminated.

We explored three classifier training methods: a Gaussian Naive Bayes (GNB), k-Nearest Neighbor (kNN), and linear Support Vector Machines (SVM). These classifiers were selected because they have been used successfully earlier by [Mitchell et al., 2003, 2004; Wang et al., 2004] for solving similar problem as well as in other applications involving high dimensional data. For example, Naive Bayes Classifiers, kNN and SVM have all been used for text classification problems [Nigam et al., 2000], where the dimensionality of the data is approximately 10^5, corresponding to the natural language vocabulary.

2.2 fMRI Task Procedure

The fMRI experiment consisted of a visuomotor sequence learning task. A set of two circles filled with different colours (Yellow, Red, Green or Blue) appeared simultaneously at two distinct positions (Top, Bottom, Left, or Right) on a screen. The correct order of the sequence is to be learnt by trial and error based on the spatial positions of the visual stimuli. In one task (P2P), the response was required to be generated according to the position of the stimuli. In another task (P2C) the subjects had to respond using an arbitrary rule with the following mapping: Green – Top, Red – Right, Blue – Bottom, and Yellow – Left. The entire sequence to be learnt consisted of six such sets. Hence this sequence learning task is called a 2 x 6 task. Subjects continue to practice the same sequence repeatedly for 4 sessions (6 sequence blocks, of 36 sec each, per session). Human brain is involved in many activities at a time. There are many background processes such as respiration, heart beat, mood, etc. that are not related to the task being performed. Hence fMRI studies must be designed to quantify relative changes in brain activity. For this purpose, the sequence blocks alternated with baseline blocks (18 sec each) in which subjects performed keypresses (using the same rule as in sequence task) viewing visual stimuli presented randomly one at a time.

2.3 Feature Selection and Data Abstraction

Data preprocessing: The case studies involve fMRI data of five normal subjects. As shape and size of brain varies for each individual, images need to be transformed into the co-ordinate system of a standard brain. This transformation would help in standardization of reporting of co-ordinates of the brain space across different studies. To achieve this, two types of brain images were collected for each subject. The first type of image referred to as a functional image, captures brain activation via the BOLD response. The second type of image, called the structural image, captures the static physical brain structure at higher resolution. First the structural images of all the subjects were standardized [Friston et al., 1995] to a template image and the resulting parameters were used for standardizing all the functional images of each subject to the template. The end result of this preprocessing step is that all the image data are in same size and approximately same shape.

Data Abstraction: As we are interested in training classifiers that work across subjects and as the strength of the BOLD response varies from subject to subject, abstracting the data is very important. Hence we adopted a normalization method in which the algorithm rescales the data in each scan of every subject such that the mean value of the resulting data is zero, i.e.
$Y^t = (X^t - meanX^t) / stdX^t,\ t = 1 \ldots n$
Where X^t's and Y^t's are the data before and after normalization, respectively and, $meanX^t$ and $stdX^t$ are the mean and the standard deviation of X^t, respectively.

The other approach for data abstraction is a method that takes the correlation of the selected features to transform the real image data.

Feature Selection: Feature selection, as a preprocessing step to machine learning, is effective in reducing dimensionality, removing irrelevant data, increasing learning accuracy, and improving result comprehensibility. However, the increase in the dimensionality of the data poses a severe challenge to many existing feature selection methods with respect to efficiency and effectiveness. The enormity of data may cause serious problems to many machine learning algorithms with respect to scalability and learning performance. For example, high dimensional data (i.e., data sets with hundreds or thousands of features) can contain high degree of irrelevant and redundant information which may greatly degrade the performance of the learning algorithms. Therefore, feature selection becomes very necessary for machine learning tasks when facing high dimensional data. The first two methods, i.e. Discrim and Active, are the same as used by [Mitchell et al., 2003, 2004; Wang et al., 2004].

Select the *n* most discriminating voxels (Discrim): In this method, voxels (features) are selected based on their ability to distinguish one target class from the other.

Select the *n* most active voxels (Active): In this method, voxels are selected based on their ability to distinguish either target class from the baseline condition.

Select the *n* feature pairs whose correlation discriminates the target classes (CorrPair): This technique is inspired from the feature selection algorithm Corona [Yang et al., 2005] which uses correlation information among features to select them. Traditional feature subset selection techniques such as Recursive Feature Elimination [Ambroise and McLachlan, 2002] and Fast Correlation Based Filter (FCBF) [Yu and Liu, 2003] lose this information as they use the feature data only. In this method, voxel pairs are selected based on their ability discriminate the target classes. The correlation value between these voxel pairs is then used as feature vector to train the classifiers.

3 Case Study

Our case study involves fMRI data collected when subjects are enagaged in visuo-motor sequence learning experiments. In these experiments, subjects learn the correct sequence of response by trial and error and continue to perform it until the end of the experiment. Therefore there are task–related changes at the neuronal level occurring continuously in different areas of the brain. These changes are reflected as changes in the activity at different voxels in fMRI images.

3.1 P2P Vs P2C Classification Study

Firstly, we have considered the problem of training classifiers to detect the cognitive states, “the subject is paying attention only towards the position of the visual stimuli” and “the subject is paying attention towards the color and the position of the visual stimuli”. When the subject is performing the P2P task he pays attention towards the position of the visual stimuli and responds only according to the position of the stimuli, the color of the stimuli is irrelevant. Whereas when the subject is performing the P2C task he pays attention towards the position of the visual stimuli and responds according to the color of the stimuli. Hence the essential difference between these two tasks is the way in which the response is generated after seeing the position of the visual stimuli.

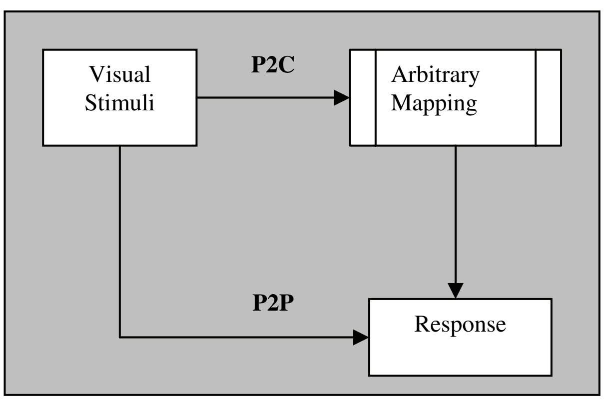

Figure 1: The Difference in the response generated during P2P and P2C Tasks

The assumption here is that since the set of brain areas underlying spatial and color perception are distinct, the cognitive states reflecting the two tasks can be distinguished based on the fMRI signals.

In this particular case study we have tried to capture this difference in our classifier function. We have used the Discriminating features and Active feature selection strategies to train a classifier function of the form f : $I_t \rightarrow$ *CognitiveState*, where I_t is an fMRI scan at a single time instance and *CognitiveState* takes values of P2P or P2C. Results of single subject classification are shown in Table 1.

For multiple subjects, the detection of instantaneous cognitive state using Discrim and Active features was below random classification. We have successfully used the correlation based feature selection to train the classifier function of the form f: $<I_1 \ldots I_t> \rightarrow$ *CognitiveState* where $<I_1 \ldots I_t>$ is a set of fMRI scans taken in a time interval of length t and *CognitiveState* takes on values P2P and P2C.

Table 1: Results of Single-Subject Classifiers detecting instantaneous cognitive states in P2P Vs P2C classification study

Feature Selection	No. of Features	GNB	KNN (k = 5)	KNN (k = 9)	SVM
Discrim	50	97.91	98.332	97.78	98.61
Discrim	100	98.19	100	99.86	98.61
Discrim	200	98.61	99.722	99.72	98.61
Active	50	92.34	98.332	97.34	98.47
Active	100	92.49	98.472	98.05	98.47
Active	200	92.91	99.026	98.61	98.47

Table 2 shows the results of using correlation based feature selection technique where the average accuracy of the multiple subject classifier is above that of a random classifier, for t taking values 12, 36 and 72. The “leave-one-subject-out” strategy was used for training multiple-subject classifiers.

Table 2: Results of Multiple subjects classifiers using CorrPair feature selection in P2P Vs P2C classification study

Feature Selection	No. of Features	GNB	KNN (k = 5)	KNN (k = 9)	SVM
t=12	200	56.66	51.66	58	51.65
t=36	200	65	67.5	65	57.5
t=72	200	75	65	75	60
t=12	474	58.33	51.66	50.8	44.1
t=36	474	75	67.5	75	60
t=72	474	80	75	80	70

All three classifiers, GNB, kNN and SVM have performed successfully and have given much better results than random-chance classification. Using the CorrPair (t=72) feature selection method gave the optimal results where the number of features is 474 and the number training samples were 16. The results show the superior performance of GNB and kNN (k=9) classifiers over the SVM.

3.2 Early Vs Late Learning Study

In the second phase, we have considered the problem of training classifiers to detect the cognitive states, “subject has learnt the sequence (Late stage where the performance is smooth and automatic)” and “subject did not learn the sequence (Early stage where the performance is deliberate and error-prone).” In both the tasks, P2P and P2C, the subjects have to learn the position sequences (or the order in which to press on the keypad, given the visual stimuli). This sequence is learnt by a process of trail and error.

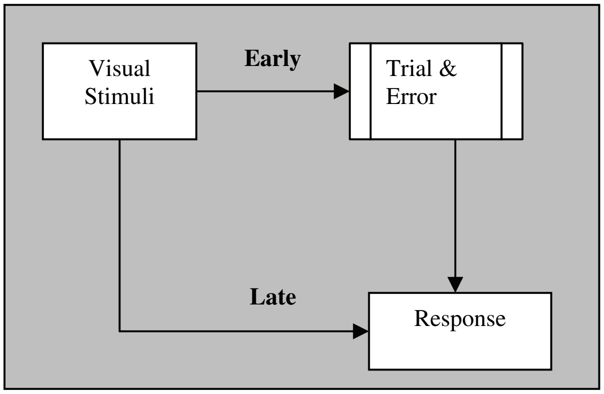

Figure 2: The Difference in the response generated during Early and Late Learning Periods

In this case study we try to detect whether the subject has learnt the sequence or not, irrespective of the task he is performing. It is to be noted that this classification task is potentially more difficult than the earlier one. The assumption here is that the brain processes related to early learning are distinguishable from those engaged in the later part. Further, it is also assumed that the cognitive states corresponding to the two learning stages are similar across the P2P and P2C tasks. We have used the Discriminating features and the Activity based feature selection methods to find a classification function of the form f : $I_t \rightarrow$ *CognitiveState* where I_t is a fMRI scan at a single time instance and *CognitiveState* takes values as either "Early Learning" or "Late Learning". As in previous case study, we were successful in training single subject classifiers.

Table 3: Results of Single-Subject Classifiers detecting instantaneous cognitive states in Early Vs Late Learning study

Feature Selection	No. of Features	GNB	KNN (k = 5)	KNN (k = 9)	SVM
Discrim	50	50	100	100	98.6
Discrim	100	60	100	100	98.6
Discrim	200	65	100	100	98.6
Active	50	80	100	100	98.6
Active	100	80	100	100	98.6
Active	200	80	100	100	98.6

We obtained less than random classification for multiple subject classifications using Discrim and Active features.
We have successfully used the correlation based feature selection to train the classifier function of the form
f: $< I_1 \ldots I_t > \rightarrow$ *CognitiveState*, where $< I_1 \ldots I_t >$ is a set of fMRI scans taken in a time interval of length t and *CognitiveState* takes on values — "Early Learning" or "Late Learning".

Table 4 shows the results of using correlation based feature selection technique where the average accuracy of the multiple subject classifier is above that of a random classifier, for t taking values 12, 36 and 72. The "leave-one-subject-out" strategy was used for training multiple-subject classifiers.

Table 4: Results of Multiple Subject classification using CorrPair feature selection in Early Vs Late: Learning Study

Feature Selection	No. of Features	GNB	KNN (k = 5)	KNN (k = 9)	SVM
t=12	200	62.5	51	62.5	52.5
t=36	200	52.5	55	65	35
t=72	200	55	50	55	30
t=12	445	56.7	50	58.3	42.5
t=36	445	62.5	43	60	50
t=72	445	55	55	50	35

All three classifiers, GNB, kNN (k=9) have performed satisfactorily and have given better results than random classification, whereas the performance of the SVM was very poor. Using the CorrPair (t=36) feature selection method gave the optimal results where the number of features is 445 and the number of training samples were 32. The results show the superior performance of GNB and kNN (k=9) classifiers over the SVM.

One important point that we have noticed is that the overall performance of the multi-subject classifier in this case study has decreased as compared to the "P2P Vs P2C" case study. One possible reason for this could be that that different subjects learn the sequence at different rates. For this purpose, we have designated the first two sessions of the experiment as the "Early Learning" period and the last two sessions as the "Late Learning" period. This criterion might not efficiently account for the subjective variations.

4 Discussion

The field of machine learning techniques for classification of data is very huge. We have tried a number of different techniques [Singh, 2005] that are suited for the high dimensional data of fMRI. The paper has taken up as a case study data from a real fMRI experiment employing visuomotor sequence learning task. Here the present state of brain is a function of its previous states and the present input (visual stimuli), a function which is unknown. Hence classification in such a dynamically varying situation is a very challenging. The learning task in this paper is inherently different from previous case studies of [Wang et al., 2004], such as the Picture Vs Sentence Study, in which the present state of the brain is just a function of its present input (picture & sentence). As with any cognitive experiments, there will be overlapping and distinct brain areas associated with the respective cognitive processes – position and colour processing in our case studies. As the focus of the present paper is detection of cognitive states, comparision of our results to the literature in cognitive neuroscience is beyond the scope.

Problems related to artifacts in data have been dealt with appropriate preprocessing techniques commonly used for FMRI data. The additional challenge posed by this data was feature selection. Hence, we have focused on comparison of different classifiers and also the feature selection strategies. The three difficulties namely, ‘high dimensionality', `different shapes & sizes of the brain' and `variations in the activity level across subjects', are the major reasons for the superior performance of the single-subject classifiers over multi-subject classifiers and are primary hurdles in the `feature selection problem'.

The existing feature selection algorithms, like FCBF and Corona, could not be implemented in such a high dimensional feature space. Hence we had to opt the strategies which were described earlier in the section 2. The drawback of these strategies is that we do not know the optimal number of features to select, and hence we have to do a lot of trial and error until we get good result or we conclude that we cannot get any better result further. In spite of these drawbacks, the correlation based strategies outperformed the activity based or discriminating feature selection methods when tested across subjects. Also, this can be successfully used in the absence of a domain expert, who identifies the regions of interest (ROI’s) in the human brain given the nature of the task and reduces the need of feature selection.

The reason for the better performance of the correlation based features as compared to the Discriminating features and the Activity based features in the multi-subject classifier training scenario is that they are not using the voxel (feature) activity values directly as feature vector to train the classifiers, instead they are abstracting the voxel values and using the correlation between the voxels as the feature vector for training, hence the problem of variations in the activity level across subjects is addressed. Whereas the Discriminating features and the Activity based features use the voxel activity values directly as features to train the classifiers and face the problem of difference in the activity levels across the subjects.

The results showed superior performance of GNB and kNN over the SVM. This was similar to Semantic Category Study of [Mitchell et al., 2003] where the number of training samples per class was very less (10). This is quiet opposite to the results obtained in the Picture versus Sentence study, where the number of training samples per class was 40 and the SVM out-performed the GNB and kNN. So we speculate that when the training samples are small in number GNB out-performs SVM.

5 Conclusions

The results obtained in the two case studies, i.e., “P2P Vs P2C Study” and “Early Vs Late Learning Study”, prove that the problem of detection of cognitive states in such a high dimensional feature space is feasible when right choice of features is made along with methods to abstract data As future work to this project we suggest the following:

1. The problem of multiple subject classifiers learning to detect the instantaneous cognitive states.
2. Finding new future selection methods for such a high dimensional and periodically changing data.

Acknowledgments

The fMRI data was collected by grants from Kawato Dynamic Brain Project, Exploratory Research for Advanced Technology, Japan. We thank Drs. Kenji Doya and Kazuyuki Samejima for their help with fMRI experiments.

References

[Ambroise and McLachlan, 2002] C. Ambroise and G. J. McLachlan. Selection bias in gene extraction on the basis of microarray gene-expression data. *Proceedings of the National Academy of Sciences*, *USA*, 99(10):6562–6566, May 2002.

[Friston et al., 1995] K. Friston, J. Ashburner, C. Frith, J. Poline, J. Heather and R. Frackowiak. Spatial registration and normalisation of images. *Human Brain Mapping*, 2:165–189, 1995.

[Mitchell, 1997] T. Mitchell. *Machine Learning.* McGraw Hill, 1997.

[Mitchell et al., 2003] T. Mitchell, R. Hutchinson, M. Just, R. Niculescu, F. Pereira and X. Wang. Classifying instantaneous cognitive states from fMRI data. In *American Medical Informatics Association Symposium*, 465–469, 2003.

[Mitchell et al., 2004] T. Mitchell, R. Hutchinson, R. Niculescu, F.Pereira, X. Wang, M. Just and S. Newman. Learning to decode cognitive states from brain images. *Machine Learning*, 57(1-2):145–175, October 2004.

[Miyapuram, 2004] K. P. Miyapuram, Visuomotor Mappings and Sequence Learning: A Whole-Brain fMRI Investigation, Master's thesis, Department of Computer and Information Sciences, University of Hyderabad, India, 2004.

[Nigam et al., 2000] K. Nigam, A. McCallum, S. Thrun and T. Mitchell. Text classification from labeled and unlabeled documents using EM. *Machine Learning*, 39:103–104, 2000.

[Ogawa et al., 1990] S. Ogawa, T. Lee, A. Kay and D. Tank, Brain magnetic resonance imaging with contrast dependent on blood oxygenation. In *Proceedings of the National Academy of Sciences, USA*, 87:9868–9872, 1990.

[Singh, 2005] Vishwajeet Singh, Detection of Cognitive States from fMRI data using Machine Learning Techniques, Master’s thesis, Department of Computer and Information Sciences, University of Hyderabad, India, 2005.

[Wang et al., 2004] X. Wang, R. Hutchinson, and T. Mitchell. Training fMRI classifiers to detect cognitive states across multiple human subjects. In *Advances in Neural Information Processing Systems*, 16:709–716, 2004.

[Yang et al., 2005] K. Yang, H. Yoon and C. Shahabi. A supervised feature subset selection technique for multivariate time series. In *Proceedings of the Workshop on Feature Selection for Data Mining: Interfacing Machine Learning with Statistics*, 92–101, 2005.

[Yu and Liu, 2003] L. Yu and H. Liu, Feature selection for high-dimensional data: A fast correlation-based filter solution. In *Proceedings of The Twentieth International Conference on Machine Leaning*, 856–863, 2003.