NeuroImage 81 (2013) 347–357

Contents lists available at SciVerse ScienceDirect

NeuroImage

journal homepage: www.elsevier.com/locate/ynimg

Multivariate decoding of brain images using ordinal regression ☆

O.M. Doyle [a,*], J. Ashburner [b], F.O. Zelaya [a], S.C.R. Williams [a], M.A. Mehta [a,1], A.F. Marquand [a,1]

[a] King's College London, Department of Neuroimaging, Institute of Psychiatry (PO89), De Crespigny Park, London SE5 8AF, UK
[b] Wellcome Trust Centre for Neuroimaging, 12 Queen Square, London WC1N 3BG, UK

ARTICLE INFO

Article history:
Accepted 3 May 2013
Available online 17 May 2013

Keywords:
Multivariate
Ordinal regression
Gaussian processes
Pharmacological MRI
Ketamine
Scopolamine

ABSTRACT

Neuroimaging data are increasingly being used to predict potential outcomes or groupings, such as clinical severity, drug dose response, and transitional illness states. In these examples, the variable (target) we want to predict is ordinal in nature. Conventional classification schemes assume that the targets are nominal and hence ignore their ranked nature, whereas parametric and/or non-parametric regression models enforce a metric notion of distance between classes. Here, we propose a novel, alternative multivariate approach that overcomes these limitations — whole brain probabilistic ordinal regression using a Gaussian process framework. We applied this technique to two data sets of pharmacological neuroimaging data from healthy volunteers. The first study was designed to investigate the effect of ketamine on brain activity and its subsequent modulation with two compounds — lamotrigine and risperidone. The second study investigates the effect of scopolamine on cerebral blood flow and its modulation using donepezil. We compared ordinal regression to multi-class classification schemes and metric regression. Considering the modulation of ketamine with lamotrigine, we found that ordinal regression significantly outperformed multi-class classification and metric regression in terms of accuracy and mean absolute error. However, for risperidone ordinal regression significantly outperformed metric regression but performed similarly to multi-class classification both in terms of accuracy and mean absolute error. For the scopolamine data set, ordinal regression was found to outperform both multi-class and metric regression techniques considering the regional cerebral blood flow in the anterior cingulate cortex. Ordinal regression was thus the only method that performed well in all cases. Our results indicate the potential of an ordinal regression approach for neuroimaging data while providing a fully probabilistic framework with elegant approaches for model selection.

© 2013 The Authors. Published by Elsevier Inc. All rights reserved.

Introduction

Neuroimaging can be used to investigate a wide range of clinical and research questions including disease processes (Fusar-Poli et al., in press; Jack et al., 2008) and decode both instantaneous cognitive states (Friston et al., 2008) and pharmacological intervention (Deakin et al., 2008). Increasingly, multivariate pattern recognition techniques are being applied to neuroimaging data to answer fundamental questions based around diagnosis/prognosis and decoding. Primarily, this can be achieved by training a learning machine using a subset of the data and their respective targets (e.g. clinical measure or state label) and observing the accuracy of the target assigned to a set of 'unseen' data, which serves to estimate the generalisation performance of the learner. Most commonly, learning machines are used to perform binary classification whereby only two labels or states are considered and the classes are assumed to have a nominal relationship to one another (e.g. patient vs. control or placebo vs. drug). To date, the most popular approach has been the binary support vector machine classifier (Fan et al., 2007; Mourao-Miranda et al., 2012; Pantazatos et al., 2012; Plant et al., 2010).

In order to consider more than two labels, multi-class learning can be employed. To perform multi-class classification a common approach, in the field of neuroimaging, has been to split the problem up into a series of binary classification problems and then apply combination strategies such as error correcting codes (Hassabis et al., 2009; Mourao-Miranda et al., 2006; Schrouff et al., 2012) or "one-versus-all" (Chu et al., 2011; Zheng et al., 2013). Alternatively, several studies have focused on the use of a single learning machine in order to consider all classes in a mutual context. Marquand et al. (2012) applied sparse multinomial logistic regression to pharmacological imaging data in order to discriminate cerebral blood flow maps collected using arterial spin labelling after placebo, atomoxetine and methylphenidate administration. Filippone et al. (2013) used an inherently multi-class Gaussian process classification approach for neuroimaging data to discriminate between different Parkinsonian neurological disorders and

☆ This is an open-access article distributed under the terms of the Creative Commons Attribution License, which permits unrestricted use, distribution, and reproduction in any medium, provided the original author and source are credited.

* Corresponding author. Fax: +44 20 3228 2116.
E-mail addresses: orla.doyle@kcl.ac.uk (O.M. Doyle), j.ashburner@ucl.ac.uk (J. Ashburner), fernando.zelaya@kcl.ac.uk (F.O. Zelaya), steve.williams@kcl.ac.uk (S.C.R. Williams), mitul.mehta@kcl.ac.uk (M.A. Mehta), andre.marquand@kcl.ac.uk (A.F. Marquand).

[1] Both authors contributed equally.

1053-8119/$ – see front matter © 2013 The Authors. Published by Elsevier Inc. All rights reserved.
http://dx.doi.org/10.1016/j.neuroimage.2013.05.036

healthy controls. Jenatton et al. (2012) investigated the performance of several sparse methods using both an inherently multi-class likelihood and a "one-versus-all" approach to discriminate the mental representation of different objects. The authors conclude that the best performance was achieved using a multinomial likelihood within their framework.

For many applications of neuroimaging, the class labels can be ordered or ranked, but it can be difficult to exactly quantify the distance between the categories. One example is a disease process continuum where the labels of scans can be ordered as: healthy controls, prodromal and disease state. The most cited example in the neuroimaging literature being the Alzheimer's Disease Neuroimaging Initiative (ADNI) data set which contains neuroimaging and clinical data from participants with Alzheimer's disease, progressive mild cognitive impairment, stable mild cognitive impairment and those who are cognitively normal (Mueller et al., 2005). Similarly, there is a lot of interest in identifying and monitoring individuals at risk of psychosis, who may be in a prodromal phase of the disease (Fusar-Poli et al., in press). For example, Borgwardt et al. (in press) analysed structural MRI scans from a psychosis continuum (healthy controls–at risk–first-episode psychosis) albeit using binary pairs of classifiers.

In addition to disease trajectories, continuum models are also evident in several other applications of neuroimaging. Increasing complexity or difficulty of cognitive tasks can be framed as a continuum; even if the task complexity can be accurately quantified (e.g. linear increase in the number of items to be remembered), the resulting increase in neural processing should not be assumed to follow the same pattern. Similarly, dose response relationships with neuroimaging data represent a continuum where the known dose interval may not match the differences in brain states (Tauscher et al., 2002). This may be particularly apparent when a small number of doses are used, or a new mechanism is being investigated. In all these examples, it would seem desirable to make use of the ordinal relationships between class labels to enable them to be more accurately predicted, compared to binary or multi-class classification approaches which ignore this information.

Here, we pose the question: how can we identify a state that is assumed to be intermediate between multiple states? We propose re-formulating this problem using a multivariate ordinal regression framework which inherently models the natural ordering in the data labels and can establish whether a collection of brain regions are ordinally related across a continuum. Crucially, this framework will consider all classes simultaneously and can be tested on a single test case (we do not require a test instance for each class to decode the ranking) in contrast to the only two previous ordinal ranking approaches for neuroimaging of which we are aware; Fan and ADNI (2011) decomposed the problem into pairwise classifiers which were then combined using an ordinal ranking rule while Pedregrosa et al. (2012) proposed training a single binary classifier to discriminate pairs of data vectors that exhaustively describe all pairs of classes. This method requires that a test case for each class must be available to decode the ranking of the image. We focus in particular on ordinal regression as conventional parametric or non-parametric regression approaches enforce a metric notion or sense of symmetry between labels. Consider the ordinal scale mild, medium and severe. To use a regression approach we may encode these labels as [1, 2, 3]. However, this metric notion assumes that the difference between mild and medium is the same as the difference between medium and severe. Additionally, unlike multi-class classification, which provides a different set of predictive weights for each class the model structure for an ordered class relation typically involves estimating a single set of predictive weights which reflects the ordering of all the classes (Chu and Ghahramani, 2005; Gutiérrez et al., 2012; Mccullagh, 1980). Moreover, performance metrics should be appropriately chosen as for ordinal regression the 'distance' of the error from its true label is of interest. That is, the magnitude of the error on incorrectly classifying mild as severe should be more highly penalised than classifying mild as moderate.

To investigate whether multivariate ordinal regression is a better suited approach for data with ordered targets we utilise two exemplar pharmacological imaging studies. The first study aims to investigate the blood oxygen level-dependent (BOLD) response to ketamine which acts as an N-methyl D-aspartate (NMDA) receptor antagonist and evokes psychotomimetic symptoms resembling schizophrenia in healthy humans (Krystal et al., 1994). Imaging markers of acute ketamine challenge have the potential to provide a powerful assay of novel therapies for psychiatric illness. In this data set, the modulation of the BOLD response to ketamine is investigated using two compounds — the anticonvulsant lamotrigine and the antipsychotic risperidone (Doyle et al., 2013). In our earlier paper, we confirmed that both lamotrigine and risperidone attenuate the effects of ketamine on the BOLD phMRI signal, albeit via different mechanisms of action (Doyle et al., 2013). We thus expect the class labels to be ordinal with placebo and ketamine at the extremities and lamotrigine or risperidone as the intermediate class. As this study was carried out to explore the level of attenuation achieved by risperidone and compare and replicate the attenuation of ketamine by lamotrigine, a priori we cannot rank lamotrigine and risperidone in a meaningful, principled manner, hence we will apply the algorithms separately. The second study focuses on the use of scopolamine which has been used for many years as a pharmacological model of 'cholinergic amnesia' gaining popularity due to the cholinergic hypothesis of geriatric memory dysfunction (Bartus et al., 1982), and the reversal of deficits is widely adopted as a tool to test putative cognitive-enhancing drugs. This study utilises arterial spin labelling to investigate the effect of scopolamine (a potent antagonist of the muscarinic acetylcholine receptor) on cerebral blood flow and its modulation using donepezil, an acetylcholinesterase inhibitor which can provide some improvement in cognitive impairments in cholinergic animal models of Alzheimer's disease as well as patients (Di Santo et al., 2013; Winblad et al., 2001). Considering that scopolamine is a non-selective muscarinic acetylcholine receptor antagonist whereas donepezil will enhance cholinergic transmission to both nicotinic and muscarinic receptors we do not expect the whole brain response to be ordinal as donepezil will produce specific regional CBF (rCBF) effects distinct from scopolamine. Therefore, we expect donepezil to attenuate scopolamine effects in regions which are rich in muscarinic receptors and have previously shown a response to scopolamine. Hence we use ordinal regression to explore the rCBF in predefined regions of interest (ROIs). Previous neuroimaging studies have shown both increases and decreases in cerebral blood flow following scopolamine administration, although these effects were not derived from fully quantitative measurements of blood flow (Grasby et al., 1995; Honer et al., 1988; Prohovnik et al., 1997). Work in experimental animals has demonstrated reduced CBF following scopolamine (Ogawa et al., 1994; Tota et al., 2012), which can be reversed by acetylcholinesterase inhibitors (Ogawa et al., 1994; Tsukada et al., 1997). Therefore, we expect rCBF to be most reduced following scopolamine administration and we expect rCBF following pre-treatment with donepezil to lie between that of placebo and scopolamine.

In this work, we describe a whole brain probabilistic approach for ordinal regression using Gaussian processes (ORGP) in a Bayesian framework. This framework provides probabilistic predictions for class membership which facilitates the quantification of uncertainty, and an elegant approach for model selection and comparison. The likelihood function of this method specifically captures the ordinal nature of the data using a threshold model which is a generalisation of the probit function as introduced by Chu and Ghahramani (2005). We will compare this approach to schemes for multi-class classification and metric regression, in all cases linear learning algorithms will be utilised. Comparisons will be carried out using three data configurations: 1) BOLD data ranked increasingly as placebo, lamotrigine–ketamine and ketamine, 2) BOLD data ranked increasingly as placebo, risperidone–ketamine and ketamine and 3) ASL data ranked

Download English Version:

https://daneshyari.com/en/article/6029123

Download Persian Version:

https://daneshyari.com/article/6029123

Daneshyari.com