A Multi-Armed Bandit to Smartly Select a Training Set from Big Medical Data

Benjamín Gutiérrez[1,2], Loïc Peter[3], Tassilo Klein[4], and Christian Wachinger[1] [*]

[1] Department of Child and Adolescent Psychiatry, Psychosomatic and
Psychotherapy, Ludwig-Maximilian-University, Munich, Germany
[2] Chair for Computer Aided Medical Procedures, Technische Universität München,
Germany
[3] Translational Imaging Group, University College London, UK
[4] SAP SE, Berlin, Germany

Abstract. With the availability of big medical image data, the selection of an adequate training set is becoming more important to address the heterogeneity of different datasets. Simply including all the data does not only incur high processing costs but can even harm the prediction. We formulate the smart and efficient selection of a training dataset from big medical image data as a multi-armed bandit problem, solved by Thompson sampling. Our method assumes that image features are not available at the time of the selection of the samples, and therefore relies only on meta information associated with the images. Our strategy simultaneously exploits data sources with high chances of yielding useful samples and explores new data regions. For our evaluation, we focus on the application of estimating the age from a brain MRI. Our results on 7,250 subjects from 10 datasets show that our approach leads to higher accuracy while only requiring a fraction of the training data.

1 Introduction

Machine learning has been one of the driving forces for the huge progress in medical imaging analysis over the last years. Of key importance for learning-based techniques is the training dataset that is used for estimating the model parameters. Traditionally, medical data has been scarce so usually all available data for a particular task was used for training. Nowadays, many initiatives make data publicly available so that huge amounts of data can potentially be used to estimate more accurate models. However, just including all the data in the training set is becoming increasingly impractical, since processing the data to create training models can be very time consuming on huge datasets. In addition, most processing may be unnecessary because it does not help the model estimation for a given task. In this work, we propose a method to select a subset of the data for training that is most relevant for a specific task. Foreshadowing some of our results, such a guided selection of a subset of the data for training

---

[*] The final publication is available on the MICCAI 17 proceedings

can lead to a higher performance than using all the available data while requiring only a fraction of the processing time.

The task of selecting a subset of data for training is challenging because at the time of making the decision, we do not yet have processed the data and we do therefore not know how the inclusion of the sample would affect the prediction. However, in many scenarios each image is assigned *metadata* about the patient (sex, diagnosis,age etc.) or the image acquisition (dataset of origin, location, imaging device, etc.). We hypothetize that some of this information can be useful to guide the selection of samples but it is a priori not clear which information is most relevant and also how it should be distributed. To address this problem, we formulate the selection of the samples to be included in a training set as reinforcement learning, where a trade-off must be reached between the exploration of new sources of data and the exploitation of sources that have been shown to lead to informative data points in the past. More specifically, we model this as a multi-armed bandit problem solved with Thompson sampling, where each arm of the bandit corresponds to a cluster of samples generated using meta information.

In this paper, we apply our sample selection method to brain age estimation [7] from MR1 T1 images. The estimated age serves as a proxy for biological age, whose difference to the chronological age can be used as indicator of disease [6,8]. The age estimation is a well-suited application for testing our algorithm as it allows us to work with a large number of datasets since the subject's age is one of the few variables that is included in every neuroimaging dataset.

1.1   Related Work

Our work is mostly related to active learning approaches, whose aim is to select samples to be labeled out of a pool of unlabeled data. Examples of active learning approaches applied to medical imaging tasks include the work by Hoi *et al.* [10], where a batch mode active learning approach was presented for selecting medical images for manually labeling the image category. Another active learning approach was proposed for the selection of histopathological slices for manual annotation in [21]. The problem was formulated as constrained submodular optimization problem and solved with a greedy algorithm. To select a diverse set of slices, the patient identity was used as meta information. From a methodological point of view, our work relates to the work of Bouneffouf *et al.* [1], where an active learning strategy based on contextual multi-armed bandits is proposed. The main difference between all these active learning approaches and our method is that image features are not available a priori in our application, and therefore can not be used in the sample selection process. Our work also relates to domain adaptation [16]. In instance weighting, the training samples are assigned weights according to the distribution of the labels (class imbalance) [11] and the distribution of the observations (covariate shift) [17]. Again these methods are not directly applicable in our scenario because not all the distribution of the metadata is defined on the target dataset.

2   Method

2.1   Incremental Sample Selection

In supervised learning, we model a predictive function $f : (\mathbf{x}, \mathbf{p}) \mapsto y$ depending on a parameter vector $\mathbf{p}$, relating an observation $\mathbf{x}$ to its label y. In our application, $\mathbf{x} \in \mathbb{R}^m$ is a vector with m quantitative brain measurements from the image and $y \in \mathbb{R}$ is the age of the subject. The parameters $\mathbf{p}$ are estimated by using a training set $S^T = \{s_1, s_2, \ldots, s_{N_{train}}\}$, where each sample $s = (\mathbf{x}, y)$ is a pair of a feature vector and its associated true label. Once the parameters are estimated, we can predict the label $\tilde{y}$ for a new observation $\tilde{\mathbf{x}}$ with $\tilde{y} = f(\tilde{\mathbf{x}}, \mathbf{p}^*)$, where the prediction depends on the estimated parameters and therefore the training dataset.

In our scenario, the samples to be included in the training set S^T are selected from a large source set $S = \{h_1, h_2, .., h_{N_{total}}\}$ containing *hidden* samples of the form $h = \{\hat{\mathbf{x}}, \hat{y}, \mathbf{m}\}$. Each h contains hidden features $\hat{\mathbf{x}}$ and label $\hat{y}$ that can only be revealed after processing the sample. In addition, each hidden sample possesses a d-dimensional vector of metadata $\mathbf{m} \in \mathbb{Z}^d$ that encodes characteristics of the patient or the image such as sex, diagnosis, and dataset of origin. In contrast to $\hat{\mathbf{x}}$ and $\hat{y}$, $\mathbf{m}$ is known a priori and can be observed at no cost. To include a sample h from set S into S^T, first its features and labels have to be revealed, which comes at a high cost. Consequently, we would like to find a sampling strategy that minimizes the cost by selecting only the most relevant samples according to the metadata $\mathbf{m}$.

2.2   Multiple Partitions of the Source Data

In order to guide our sample selection algorithm, we create multiple partitions of the source dataset, where each one considers different information from the metadata $\mathbf{m}$. Considering the j-th meta information ($1 \leq j \leq d$), we create the j-th partition $S = \cup_{i=1}^{\eta_j} C_i^j$ with η_j a predefined number of bins for $\mathbf{m}[j]$. As a concrete example, sex could be used for partitioning the data, so $S = C_{\text{female}}^{\text{sex}} \cup C_{\text{male}}^{\text{sex}}$ and $\eta_{\text{sex}} = 2$. All the clusters generated using different meta information are merged into a set of clusters $\mathcal{C} = \{C_i^j\}$. We hypothesize that given this partitioning, there exist clusters $C_i \in \mathcal{C}$ that contain more relevant samples than others for a specific task. Intuitively, we would like to draw samples h from clusters with a higher probability of returning a relevant sample. However, since the relationship between the metadata and the task is uncertain, the utility of each cluster for a specific task is unknown beforehand. We will now describe a strategy that simultaneously *explores* the clusters to find out which ones contain more relevant information and *exploits* them by extracting as many samples from relevant clusters as possible.

2.3   Sample selection as a multi-armed bandit problem

We model the task of sequential sample selection as a multi-armed bandit problem. At each iteration t, a new sample is added to the training dataset S^T. For

adding a sample, the algorithm decides which cluster $C_i \in \mathcal{C}$ to exploit and randomly draws a training sample s_t from cluster C_i. The corresponding feature vector $\mathbf{x}_t$ and label y_t are revealed and the usefulness of the sample s_t for the given task is evaluated, yielding a reward $r_t \in \mathbb{R}$. A reward $r_t = 1$ is given if adding the sample improves the prediction accuracy of the model and $r_t = -1$ otherwise.

At $t = 0$, we do not possess knowledge about the utility of any cluster. However, this knowledge is incrementally built as more and more samples are drawn and their rewards are revealed. To this end, each cluster is assigned a distribution of rewards Π_i. With every sample the distribution better approximates the true expected reward of the cluster, but every new sample also incurs a cost. Therefore, a strategy needs to be designed that explores the distribution for each of the clusters, while at the same time exploiting as often as possible the most rewarding sources.

To solve the problem of selecting from which C_i to sample at every iteration t, we follow a strategy based on Thompson sampling [18] with binary rewards. In this setting, the expected rewards are modeled using a probability P_i following a Bernoulli distribution with parameter $\pi_i \in [0, 1]$. We maintain an estimate of the likelihood of each π_i given the number of successes α_i and failures β_i observed for the cluster C_i so far. Successes $(r = 1)$ and failures $(r = -1)$ are defined based on the reward of the current iteration. It can be shown that this likelihood follows the conjugate distribution of a Bernoulli law, i.e., a Beta distribution $Beta(\alpha_i, \beta_i)$ so that

$$P(\pi_i | \alpha_i, \beta_i) = \frac{\Gamma(\alpha_i + \beta_i)}{\Gamma(\alpha_i)\Gamma(\beta_i)}(1 - \pi_i)^{\beta_i - 1}\pi_i^{\alpha_i - 1}. \tag{1}$$

with the gamma function Γ. At each iteration, $\hat{\pi}_i$ is drawn from each cluster distribution P_i and the cluster with the maximum $\hat{\pi}_i$ is chosen. The procedure is summarized in Algorithm 1.

Algorithm 1 Thompson Sampling for Sample Selection

1: $\alpha_i = 1, \beta_i = 1, \forall i \in \{1, \ldots, N\}$
2: **for** $t = 1, 2, \ldots$ **do**
3:     **for** $i = 1, \ldots, N$ **do**
4:         Draw $\hat{\pi}_i$ from $Beta(\alpha_i, \beta_i)$.
5:     Reveal sample $h_t = \{\mathbf{x}_t, y_t, \mathbf{m}_t\}$ from cluster C_j where $j := \arg\max_i \hat{\pi}_i$.
6:     Add sample h_t to S^T and remove from all clusters.
7:     Obtain new model parameters $\mathbf{p}^*$ from updated training set S^T.
8:     Compute reward r_t based on new prediction $\tilde{y} = f(\mathbf{x}, \mathbf{p}^*)$.
9:     **if** $r_t == 1$ **then** $\alpha_j = \alpha_j + 1$
10:     **else** $\beta_j = \beta_j + 1$

---

3   Results

In order to showcase the advantages of our multi-armed bandit sampling algorithm (MABS), we present an evaluation of our method in the task of estimating the biological age of a subject given a set of volume and thickness features of the brain. We choose this task in particular because of the big number of available brain images in public databases and, the relevance of age estimation as a tool for diagnostic of neuro degenerative diseases [8,19]. For predicting the age, we reconstruct brain scans with FreeSurfer [5] and extract volume and thickness measurements to create our feature vectors $\mathbf{x}$. Based on these features, we train a regression model for predicting the age of previously unseen subjects.

3.1   Data

We work on MRI T1 brain scans from 10 large-scale public datasets: ABIDE [3], ADHD200 [15], AIBL [4], COBRE [14], IXI[5], GSP [2], HCP [20], MCIC [9], PPMI [13] and OASIS [12]. From all these datasets we obtain a total number of 7,250 images, which is to the best of our knowledge, the largest dataset collected for the task of age prediction. Since each one of these datasets is targeted towards different applications, the selected population is heterogeneous in terms of age, sex, and health status. Images are processed with FreeSurfer [5] and thickness and volume measurements extracted. Even though this is a fully automatic tool, the extraction of the feature is a computationally intensive task which is by far the bottleneck of our age prediction regression model.

3.2   Age estimation

We perform age estimation on two different testing scenarios. In the first, we create a testing dataset by randomly selecting subsets from all the datasets. The aim of this experiment is to show that our method is capable of selecting samples that will create a model that can generalize well to a heterogeneous population. In the second scenario, the testing dataset corresponds to a single dataset. In this scenario, we show that the sample selection permits tailoring the training dataset to a specific target dataset.

Experiment 1. For the first experiment we take all the images in the dataset and we divide them randomly into three sets: 1) a small validation set of 2% of all samples to compute the rewards given to MABS , 2) a large testing set of 48% to measure the performance of our age regression task, and 3) a large hidden training set of 50%, from which samples are taken sequentially using MABS. We perform the sequential sample selection described in Algorithm 1 using the following metadata to construct the clusters $\mathcal{C}$: *age*, *dataset*, *diagnosis*, and *sex*. We experiment with considering all of the metadata separately, to investigate the importance of each one, and the joint modeling. Since we require to evaluate a regression model every time a sample is included, we opted to use ridge regression

---

[5] http://brain-development.org/ixi-dataset/

as our learning algorithm. Rewards r are given to each bandit estimating and observing if the r^2 score of the prediction in the validation set increases. It is important to emphasize that the testing set is not observed by the bandits in the process of giving rewards. Every experiment is repeated 20 times using different random splits and the mean results are shown. We compare with two baselines: the first one (RANDOM) consists of obtaining samples at random from the hidden set and adding them sequentially to the training set. As a second baseline (AGE PRIOR), we add samples sequentially by following the age distribution of the testing set . The results of this first experiment are shown in figure 1 (top left). In almost all the cases, using MABS as a selection strategy performed better than the baselines. It is important to observe that an increase in performance is obtained not only when the relationships between the metadata and the task are direct, like in the case of the clusters constructed by age, but also when this relationship is not clear like in the case of clustering the images using only dataset or diagnostic information. Another important aspect is that even when the meta information is not informative, like in the case of the clusters generated by sex, the prediction using MABS is not affected.

Experiment 2. For our second experiment, we perform age estimation with the test data being a specific dataset. This experiment follows the same methodology as the previous one with the important difference of how the datasets are split. This time the split is done by choosing: 1) a small validation set, taken only from the target dataset, 2) a testing set, which corresponds to the remaining samples in the target dataset not included in the validation set, and 3) a hidden dataset containing all the samples from the remaining datasets. The goal of this experiment is to show that our methodology can be deployed when samples have to be selected according to a specific population and prediction task. From the results in figure 1, we observe that depending on the dataset, using bandits with only one specific source of meta information can actually improve the sample selection algorithm. However, the best meta information for a particular task is different in every case. We also observe that in general the MABS using all available meta information extracts informative samples more efficiently than our baselines and always close to the best performing single meta information MABS. This reinforces our hypothesis that it is hard to define an a priori relationship between the meta data and the task, and is therefore a better strategy to let MABS select from multiple sources of meta information at once.

4   Conclusion

We have proposed a method for efficiently and intelligently sampling a training dataset from a large pool of data. The problem was formulated as reinforcement learning, where the training dataset was sequentially built after evaluating a reward function at every step. Concretely, we used a multi-armed bandit model that was solved with Thompson sampling. The intelligent selection considered metadata of the scan to construct a distribution about the expected reward of a training sample. Our results showed that the selective sampling approach
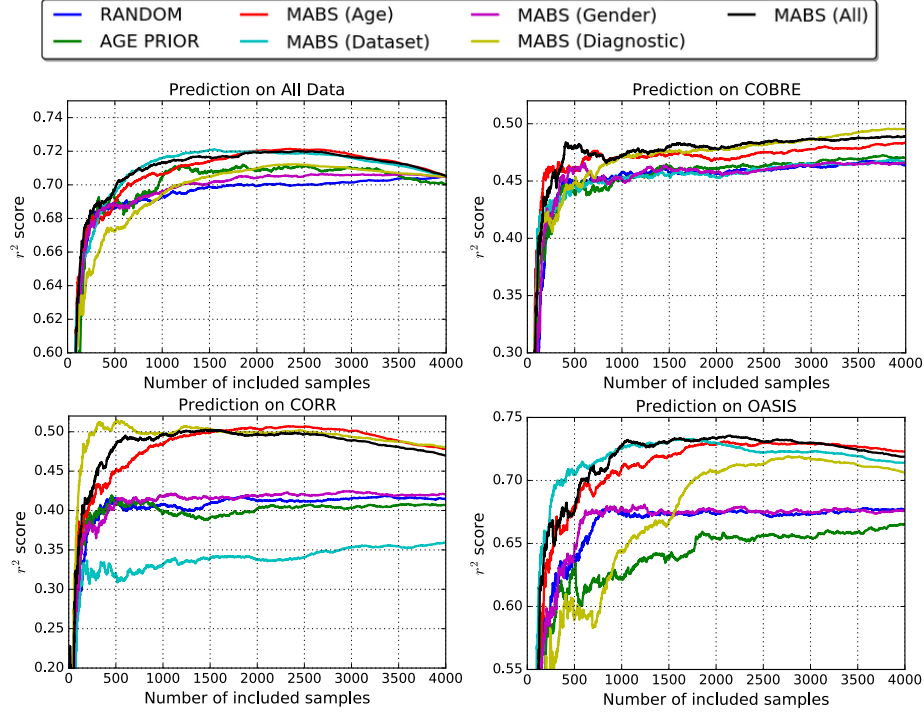
Fig. 1. Results of our age prediction experiments in terms of r^2 score. A comparison is made between MABS using different strategies to build the clusters $\mathcal{C}$, a random selection of samples, and a random selection based on the age distribution of the test data. To improve the visualization of the results, we limit the plot to 4,000 samples.

leads to higher accuracy than using all the data, while requiring less time for processing the data. We demonstrated that our technique can either be used to build a general model or to adapt to a specific target dataset, depending on the composition of the test dataset. Since our method does not require to observe the information contained in the images, it can also be applied to predict useful samples even before the images are acquired, guiding the recruitment of subjects.

5   Acknowledgements

References

1. D. Bouneffouf, R. Laroche, T. Urvoy, R. Féraud, and R. Allesiardo. Contextual bandit for active learning: Active thompson sampling. In *International Conference on Neural Information Processing*, pages 405–412. Springer, 2014.

2. R. Buckner, M. Hollinshead, A. Holmes, D. Brohawn, J. Fagerness, T. O'Keefe, and J. Roffman. The brain genomics superstruct project. *Harvard Dataverse Network*, 2012.

3. A. Di Martino, C.-G. Yan, Q. Li, E. Denio, F. X. Castellanos, K. Alaerts, J. S. Anderson, M. Assaf, S. Y. Bookheimer, M. Dapretto, et al. The autism brain imaging data exchange: towards a large-scale evaluation of the intrinsic brain architecture in autism. *Molecular psychiatry*, 19(6):659–667, 2014.

4. K. A. Ellis, A. I. Bush, D. Darby, D. De Fazio, J. Foster, P. Hudson, N. T. Lautenschlager, N. Lenzo, R. N. Martins, P. Maruff, et al. The australian imaging, biomarkers and lifestyle (aibl) study of aging: methodology and baseline characteristics of 1112 individuals recruited for a longitudinal study of alzheimer's disease. *International Psychogeriatrics*, 21(04):672–687, 2009.

5. B. Fischl, D. H. Salat, E. Busa, M. Albert, M. Dieterich, C. Haselgrove, A. van der Kouwe, R. Killiany, D. Kennedy, S. Klaveness, A. Montillo, N. Makris, B. Rosen, and A. M. Dale. Whole brain segmentation: automated labeling of neuroanatomical structures in the human brain. *Neuron*, 33(3):341–355, 2002.

6. K. Franke, E. Luders, A. May, M. Wilke, and C. Gaser. Brain maturation: predicting individual brainage in children and adolescents using structural mri. *Neuroimage*, 63(3):1305–1312, 2012.

7. K. Franke, G. Ziegler, S. Klöppel, C. Gaser, A. D. N. Initiative, et al. Estimating the age of healthy subjects from t 1-weighted mri scans using kernel methods: Exploring the influence of various parameters. *Neuroimage*, 50(3):883–892, 2010.

8. C. Gaser, K. Franke, S. Klöppel, N. Koutsouleris, H. Sauer, A. D. N. Initiative, et al. Brainage in mild cognitive impaired patients: predicting the conversion to alzheimers disease. *PloS one*, 8(6):e67346, 2013.

9. R. L. Gollub, J. M. Shoemaker, M. D. King, T. White, S. Ehrlich, S. R. Sponheim, V. P. Clark, J. A. Turner, B. A. Mueller, V. Magnotta, et al. The mcic collection: a shared repository of multi-modal, multi-site brain image data from a clinical investigation of schizophrenia. *Neuroinformatics*, 11(3):367–388, 2013.

10. S. C. Hoi, R. Jin, J. Zhu, and M. R. Lyu. Batch mode active learning and its application to medical image classification. In *Proceedings of the 23rd international conference on Machine learning*, pages 417–424. ACM, 2006.

11. N. Japkowicz and S. Stephen. The class imbalance problem: A systematic study. *Intelligent data analysis*, 6(5):429–449, 2002.

12. D. S. Marcus, T. H. Wang, J. Parker, J. G. Csernansky, J. C. Morris, and R. L. Buckner. Open access series of imaging studies (oasis): cross-sectional mri data in young, middle aged, nondemented, and demented older adults. *Journal of cognitive neuroscience*, 19(9):1498–1507, 2007.

13. K. Marek, D. Jennings, S. Lasch, A. Siderowf, C. Tanner, T. Simuni, C. Coffey, K. Kieburtz, E. Flagg, S. Chowdhury, et al. The parkinson progression marker initiative (ppmi). *Progress in neurobiology*, 95(4):629–635, 2011.

14. A. R. Mayer, D. Ruhl, F. Merideth, J. Ling, F. M. Hanlon, J. Bustillo, and J. Cañive. Functional imaging of the hemodynamic sensory gating response in schizophrenia. *Human brain mapping*, 34(9):2302–2312, 2013.

15. M. P. Milham, D. Fair, M. Mennes, S. H. Mostofsky, et al. The adhd-200 consortium: a model to advance the translational potential of neuroimaging in clinical neuroscience. *Frontiers in systems neuroscience*, 6:62, 2012.
16. S. J. Pan and Q. Yang. A survey on transfer learning. *Knowledge and Data Engineering, IEEE Transactions on*, 22(10):1345–1359, 2010.
17. H. Shimodaira. Improving predictive inference under covariate shift by weighting the log-likelihood function. *Journal of statistical planning and inference*, 90(2):227–244, 2000.
18. W. R. Thompson. On the likelihood that one unknown probability exceeds another in view of the evidence of two samples. *Biometrika*, 25(3/4):285–294, 1933.
19. S. Valizadeh, J. Hänggi, S. Mérillat, and L. Jäncke. Age prediction on the basis of brain anatomical measures. *Human Brain Mapping*, 38(2):997–1008, 2017.
20. D. C. Van Essen, S. M. Smith, D. M. Barch, T. E. Behrens, E. Yacoub, K. Ugurbil, W.-M. H. Consortium, et al. The wu-minn human connectome project: an overview. *Neuroimage*, 80:62–79, 2013.
21. Y. Zhu, S. Zhang, W. Liu, and D. N. Metaxas. Scalable histopathological image analysis via active learning. In *International Conference on Medical Image Computing and Computer-Assisted Intervention*, pages 369–376. Springer, 2014.