Ask the doctor – Improving drug sensitivity predictions through active expert knowledge elicitation

Iiris Sundin^{1,*}, Tomi Peltola¹, Muntasir Mamun Majumder², Pedram Daee¹, Marta Soare¹, Homayun Afrabandpey¹, Caroline Heckman², Samuel Kaski^{1,†,*} and Pekka Marttinen^{1,†,*}

¹Helsinki Institute for Information Technology HIIT, Department of Computer Science, Aalto University, Finland ²Institute for Molecular Medicine Finland (FIMM), University of Helsinki, Finland.

Abstract

Predicting the efficacy of a drug for a given individual, using high-dimensional genomic measurements, is at the core of precision medicine. However, identifying features on which to base the predictions remains a challenge, especially when the sample size is small. Incorporating expert knowledge offers a promising alternative to improve a prediction model, but collecting such knowledge is laborious to the expert if the number of candidate features is very large. We introduce a probabilistic model that can incorporate expert feedback about the impact of genomic measurements on the sensitivity of a cancer cell for a given drug. We also present two methods to intelligently collect this feedback from the expert, using experimental design and multi-armed bandit models. In a multiple myeloma blood cancer data set (n=51), expert knowledge decreased the prediction error by 8%. Furthermore, the intelligent approaches can be used to reduce the workload of feedback collection to less than 30% on average, compared to a naive approach.

1 Introduction

In genomics-based precision medicine, prediction is challenging due to large-scale -omics data with small sample size. For example, in cancer studies the sample size is typically less than a thousand of cell lines, and even fewer patients, whereas high-throughput methods produce thousands of genomic and molecular features for each observation.

A natural way to deal with the problems caused by a small sample size is to measure more data. This is, however, often not an available option, due to costs, risks, or the rarity of the disease. A more rarely exploited alternative is to ask an expert. *Prior elicitation techniques* [1] have been used in Bayesian data analysis for constructing prior distributions that take into account expert knowledge, and hence can restrict the range of parameters to be later used in learning models [2, 3, 4]. These techniques focus on how to reliably elicit knowledge, whereas in practice it is equally important to minimize the effort required from the expert. *Interactive and sequential learning* can help by carefully deciding what to ask the expert [5, 6, 7]. In this work, we leverage on the recent initial work on interactive expert knowledge elicitation [8, 9, 10], and introduce sequential knowledge elicitation methods to the precision medicine prediction task, illustrated in Figure 1. As a challenging case study, we predict drug responses of *ex vivo* cell samples from blood cancer patients (n = 51), based on mutation data and cytogenetics markers (in total 3032 features). Two well-informed experts were asked to provide feedback about the relevance of features when predicting sensitivity of a cell-line to specific targeted drugs (*relevance* feedback). In

^{*}To whom correspondence should be addressed.

[†]Equally contributing PIs.



Figure 1: Overview. Predictions in small-sample-size problems are improved by asking experts in an elicitation loop. The system presents questions about relevance of the features to the expert sequentially, in order to maximize performance with a minimal number of questions, i.e., on a budget. The expert answers the questions by indicating the relevance of a feature and whether it is positively or negatively correlated with the drug response.

addition, we give the experts the option of telling if a feature is positively or negatively correlated with the drug response (*directional* feedback).

Our main contribution is to show, for the first time, that sequential expert knowledge elicitation can improve predictive modeling with high-throughput omics data in precision medicine.

2 Models and algorithms

2.1 Prediction model

A sparse linear regression model is used to predict drug sensitivities based on genomic features and elicited expert knowledge. Let $y_{n,d}$ be the sensitivity of the *n*th patient for drug *d*, and $\boldsymbol{x}_n \in \mathbb{R}^M$ be the vector of the patient's *M* genomic features. We assume a Gaussian observation model $y_{n,d} \sim N(\boldsymbol{w}_d^\top \boldsymbol{x}_n, \sigma_d^2)$, where the $\boldsymbol{w}_d \in \mathbb{R}^M$ are the regression weights and σ_d^2 is the residual variance. We use a sparsity-inducing spike-and-slab prior on the regression weights [11, 12]. Expert knowledge is incorporated into the model via feedback observation models [9]. We extend the work from [9] to include directional feedback.

2.2 Expert knowledge elicitation methods

The purpose of expert knowledge elicitation algorithms is to sequentially choose queries to the expert, so that the improvement in predictions is maximized.

Sequential experimental design. We introduce a sequential experimental design approach to select the next (drug, feature) pair candidate to be queried for feedback from the expert, extending the work in [9]. At each iteration, we find the pair where the feedback from the expert is expected to have the maximal influence on the drug sensitivity prediction. The amount of information in the expert feedback is measured by the Kullback–Leibler divergence (KL) between the predictive distributions before and after observing the feedback. As the feedback value itself is unobserved before the actual query, an expectation over the predictive distributions of the two types of feedbacks is computed in finding the (drug, feature) pair with the highest expected information gain.

User model. We introduce an alternative additional approach for selecting the next (drug, feature) pair candidate using a multi-armed bandit user model. We borrow this idea from the bandit literature (see, for instance, [13]) to ensure that our user model concentrates the queries to the (drug, feature) pairs that are likely to get an answer from the expert. We extend the work in [10] by using biological prior information from DrugBank [14] and KEGG pathways in Molecular Signatures Database (MSigDB) [15] to describe candidate pairs.



Figure 2: Performance improves faster with the active elicitation methods than with randomly selected feedback queries. The curves show mean squared errors as a function of the number of iterations for the three query methods, with feedback of the senior researcher (left) and doctoral candidate (right). In each iteration, relevance of a (drug, feature) pair is queried from the expert. The 50 independent runs of randomly selected queries are shown in red, and their average in thick red line.

3 Experimental results

In order to evaluate the proposed methods, we applied them to real patient data and used feedback from well-informed experts¹. We simulated sequential expert knowledge elicitation by iteratively querying (drug, feature) pairs for feedback, and answering the queries using the pre-collected feedback. We present here the two main results of the experiments.

Expert knowledge elicitation improves the accuracy of drug sensitivity prediction. Table 1 establishes the baselines and shows that our model has performance comparable to the standard prediction models² in leave-one-out cross-validation, without expert feedback. The main result is that feedback of both of the experts improves the predictions, as can be seen in Table 1. The model with full feedback from the senior researcher has 7% higher C-index and 8% lower MSE compared to the no-feedback model, and is confidently better (bootstrapping over the predictions for the samples gives probabilities 0.98 for C-index and 0.95 for MSE of the model with feedback performing better than the no-feedback model).

Table 1: Performance of drug sensitivity prediction without expert feedback in baseline models and our spikeand-slab regression model. Comparison to the performance of our spike-and-slab regression model with full expert feedback from SR = Senior Researcher, and DC = Doctoral Candidate. Values are averaged over the 12 drugs. Best results with and without feedback on each row have been boldfaced.

	Without feedback				With full feedback	
	Data mean	Ridge	Elastic net	Spike-and-slab	SR	DC
C-index	0.50	0.62	0.60	0.61	0.65	0.63
MSE	1.06	0.94	1.00	0.93	0.86	0.92

Sequential knowledge elicitation reduces the number of queries required from the expert. Figure 2 shows that the sequential knowledge elicitation methods achieve lower error in prediction accuracy with fewer feedbacks than random selection. On average, the sequential experimental design requires only 23% of the number of queries compared to random, and the bandit user model 32%, to achieve half of the potential improvement.

¹Collected information: We asked questions from two well-informed experts of multiple myeloma, using a form containing 161 mutations known to be related to cancer [16], and 7 cytogenetic markers. The experts were asked to evaluate relevance of a feature to the response of 12 targeted drugs, grouped by the targets (BCL-2, Glucocorticoid, PI3K/mTOR, and MEK1/2).

 $^{^{2}}$ Ridge regression and elastic net are implemented using the glmnet R-package [17] with nested cross-validation for choosing the regularization parameters.

4 Conclusion

In this extended abstract, we report the work where we showed, for the first time, that sequential expert knowledge elicitation improves drug sensitivity prediction in precision cancer medicine. We also showed, in a simulated user experiment with real expert feedback, that the proposed algorithms can elicit knowledge from experts efficiently. The results indicate that expert knowledge can be very beneficial and, hence, should be taken into account in modeling tasks of precision medicine. We found that the most efficient elicitation method was different for the two experts. An obvious next question is how to combine the two elicitation methods to optimally utilize the complementary principles in them. In the future we will carry out a wider study to thoroughly quantify the effect of expert feedback, and to investigate further the initial observations about the impact of the type of feedback and the level of seniority of the experts.

Acknowledgements

This work was supported by the Academy of Finland [grant numbers 295503, 294238, 292334, 286607, 294015] and Centre of Excellence in Computational Inference Research COIN; and by Jenny and Antti Wihuri Foundation. We acknowledge the computational resources provided by the Aalto Science-IT project.

References

- A. O'Hagan, C. E. Buck, A. Daneshkhah, J. R. Eiser, P. H. Garthwaite, D. J. Jenkinson, J. E. Oakley, and T. Rakow, Uncertain Judgements: Eliciting Experts' Probabilities. Chichester, England: Wiley, 2006.
- [2] P. H. Garthwaite, S. A. Al-Awadhi, F. G. Elfadaly, and D. J. Jenkinson, "Prior distribution elicitation for generalized linear and piecewise-linear models," *Journal of Applied Statistics*, vol. 40, no. 1, pp. 59–75, 2013.
- [3] J. B. Kadane, J. M. Dickey, R. L. Winkler, W. S. Smith, and S. C. Peters, "Interactive elicitation of opinion for a normal linear model," *Journal of the American Statistical Association*, vol. 75, no. 372, pp. 845–854, 1980.
- [4] H. Afrabandpey, T. Peltola, and S. Kaski, "Interactive prior elicitation of feature similarities for small sample size prediction," in *Proceedings of the 25th International Conference on User Modelling, Adaptation and Personalization* (UMAP '17) (To appear), arXiv preprint arXiv:1612.02802, 2017.
- [5] Z. Lu and T. K. Leen, "Semi-supervised clustering with pairwise constraints: A discriminative approach," in Proc of AISTATS, pp. 299–306, 2007.
- [6] M.-F. Balcan and A. Blum, "Clustering with interactive feedback," in International Conference on Algorithmic Learning Theory, pp. 316–328, Springer, 2008.
- [7] L. House, L. Scotland, and C. Han, "Bayesian visual analytics: BaVa," Statistical Analysis and Data Mining, vol. 8, no. 1, pp. 1–13, 2015.
- [8] M. Soare, M. Ammad-ud-din, and S. Kaski, "Regression with n → 1 by Expert Knowledge Elicitation," in Proceedings of the 15th IEEE ICMLA International Conference on Machine learning and Applications, pp. 734–739, 2016.
- [9] P. Daee, T. Peltola, M. Soare, and S. Kaski, "Knowledge elicitation via sequential probabilistic inference for highdimensional prediction," in arXiv preprint arXiv:1612.03328, 2016.
- [10] L. Micallef, I. Sundin, P. Marttinen, M. Ammad-ud-din, T. Peltola, M. Soare, G. Jacucci, and S. Kaski, "Interactive elicitation of knowledge on feature relevance improves predictions in small data sets," in *IUI2017, to appear*, 2017.
- [11] T. J. Mitchell and J. J. Beauchamp, "Bayesian variable selection in linear regression," Journal of the American Statistical Association, vol. 83, no. 404, pp. 1023–1032, 1988.
- [12] E. I. George and R. E. McCulloch, "Variable selection via Gibbs sampling," Journal of the American Statistical Association, vol. 88, no. 423, pp. 881–889, 1993.
- T. L. Lai and H. Robbins, "Asymptotically efficient adaptive allocation rules," Advances in Applied Mathematics, vol. 6, no. 1, pp. 4–22, 1985.
- [14] D. S. Wishart, C. Knox, A. C. Guo, S. Shrivastava, M. Hassanali, P. Stothard, Z. Chang, and J. Woolsey, "DrugBank: a comprehensive resource for in silico drug discovery and exploration," *Nucleic acids research*, vol. 34, pp. D668–D672, 2006.
- [15] A. Liberzon, A. Subramanian, R. Pinchback, H. Thorvaldsdóttir, P. Tamayo, and J. P. Mesirov, "Molecular signatures database (MSigDB) 3.0," *Bioinformatics*, vol. 27, no. 12, pp. 1739–1740, 2011.
- [16] S. A. Forbes, D. Beare, P. Gunasekaran, K. Leung, N. Bindal, H. Boutselakis, M. Ding, S. Bamford, C. Cole, S. Ward, C. Y. Kok, M. Jia, T. De, J. W. Teague, M. R. Stratton, U. McDermott, and P. J. Campbell, "COSMIC: exploring the world's knowledge of somatic mutations in human cancer," *Nucleic Acids Research*, vol. 43, pp. D805–D811, 2014.
- [17] J. Friedman, T. Hastie, and R. Tibshirani, "Regularization paths for generalized linear models via coordinate descent," *Journal of Statistical Software*, vol. 33, no. 1, pp. 1–22, 2010.