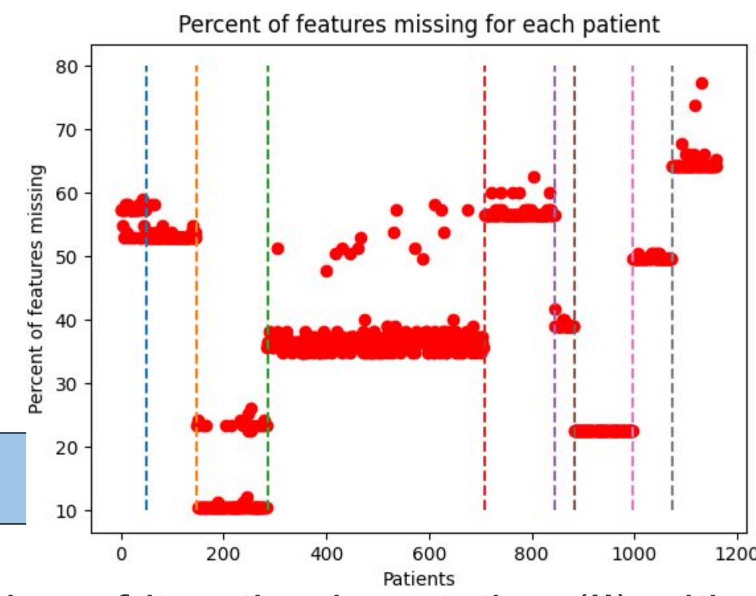
Predicting anxiety treatment outcomes with machine learning

Marija Stanojevic, Lesley A. Norris, Philip C. Kendall, Zoran Obradovic
Temple University, Philadelphia, USA



Introduction

- **Background**:
  - Anxiety is the most common disorder among adolescents [1].
  - Anxiety or depression in adolescents from 5.4% in 2003 to 8.4% in 2012. [2]
- **Objective**: Develop machine learning tool to facilitate evidence-based anxiety treatment allocation decisions.
- **Challenges**:
  - Data comes from multi-informant, randomized, longitudinal clinical trials.
  - Studies have more features than patients, even after clinical feature selection.
  - The number of patients who withdraw from treatment or do not fill in all the information is significant.
  - Diagnostic comorbidity is the norm, leading to multiple predictive variables.
  - Data is influenced by recorded diagnoses, physicians who conducted studies, and geographical regions.

Experiments

- **Data:**
  - From 9 studies on youth (1,161 samples) with primary anxiety disorders.
  - Diagnoses range is 0-8, with higher numbers indicating greater severity.
  - 108 features selected on the basis of domain knowledge
  - 202 youth who dropped early were used to understand the properties of domain adaptation to withdrawn patients.
- **Baseline methods:**
  - 1) Statistics (mean, median, mode, random, MCMC);
  - 2) Nearest neighbors (KNN, FKM, KI, FCKI, MICE);
  - 3) Decision trees (DT) and random forests (RF) (missForest);
  - 4) Matrix decomposition (soft imputation, singular value decomposition).
- **Iterative imputation with regressors**:
  - 1) Simple regressions (linear, Bayesian ridge (BR), orthogonal matching pursuit (OMP), Bayesian automatic relevance determination (ARD))
  - 2) KNN regression
  - 3) Ensambles (gradient boosting (GB) with DT, adaboost with DT or elastic nets (EN), bag of DT or EN).
- **Imputation evaluation**: 10% of non-missing values are randomly masked. Then, imputation algorithms are trained and evaluated on that dataset. Root mean square error (RMSE) is calculated on masked data by comparing original and imputed values (Full RMSE). When imputation is learned and performed on each study separately, the datasets are joined before RMSE is calculated on all data at once (Avg study RMSE). We also calculate imputation RMSE for data from each study separately.
- **Prediction evaluation:** Data is split ratio of 70:15:15. The features are normalized. Prediction is evaluated using RMSE on non-imputed predicted values of validation and test data to avoid underestimation due to imputation of predicted variables.

Conclusions

- Dataset's missing rate, size, and structure of missing values influence the ability to predict youth anxiety well.
- Choosing the proposed iterative imputation with the bagging of elastic net regressions gives a slight advantage despite the chosen predictor. Joining datasets is more significant in successful prediction, especially with the out-of-domain target.
- There is a vast gap in data and method understanding and software availability. While our paper gives the first directions on handling youth anxiety data with machine learning, additional work is required to achieve better results on this data.

Acknowledgement
We thank D. H. Bodden, S. M. Bogels, M. H. Nauta, E. De Haan, J. Ringrose, C. Appelboom, A. G. Brinkman, K. C. M. M. J. Appelboom-Geerts, E. Flannery-Schroeder, S. M. Panichelli-Mindel, M. Southam-Gerow, A. Henin, M. Warman, J. L. Hudson, E. Gosch, C. Suveg, A. Scholing, P. M. Emmelkamp, R. B. Minderaa, J. T. Walkup, A. M. Albano, J. Piacentini, B. Birmaher, S. N. Compton, J. T. Sherrill, G. S. Ginsburg, M. A. Rynn, J. McCracken, B. Waslick, S. Iyengar, J. S. March, J. J. Wood, M. Southam-Gerow, B. C. Chu, M. Sigman, M. A. Villabø, M. Narayanan, S. N. Compton, S. P. Neumer, K. Aafjes-van Doorn, C. Kamsteeg, J. Bate and M. Aafjes for providing the anonymized data from their studies cited in the paper.

Methods

- **Imputation methods**: We propose the integration of iterative imputation (II) with ensemble methods in two-level iterative imputation:
  - The outer level iteration step finishes after all features are imputed within that step. Some of the relevant features have a missing rate of 80% (**Figure 1**).
  - The inner iteration step always handles the feature with the lowest missing rate among the features not handled in the current outer iteration step.
  - We train separate imputing models on X and y variables.
- **Prediction methods**: We trained and tested diverse regression models to understand relationship between imputation and prediction models.
- **Joining multiple studies:** We propose combining studies into a single dataset to achieve better prediction confidence.
- **Domain adaptation:** Data of withdrawn patients have even higher missing rate and an additional bias. The proposed imputation approach using joined data from multiple studies can improve transfer learning prediction for the case of withdrawn patients.

Results

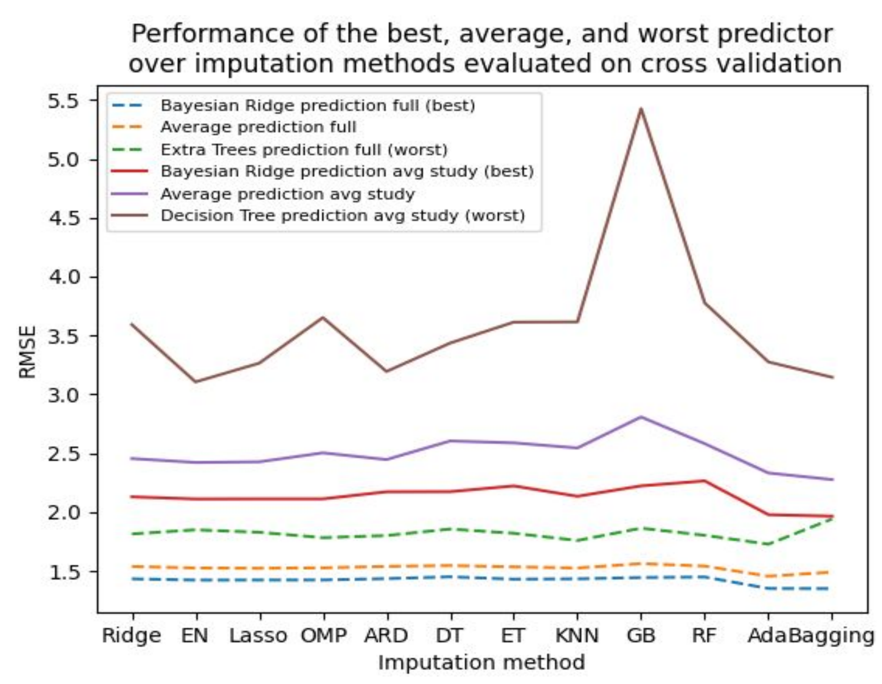
- **Imputation results: Table I** shows results when 1) joined dataset is imputed (Full RMSE) and 2) each study is imputed separately and results are combined (Avg study RMSE). The first part shows baseline methods, followed by the proposed methods. Six proposed methods have better Full RMSE than the best baseline, and a bag of elastic nets performs the best among all.
- **Performance comparison between joined and single datasets:** Using joined dataset increases RMSE performance by 9.3% for the best imputation model (II with a bag of elastic net regressions). It decreases the variability of imputation performance. **Table II** shows RMSE of imputation per each study. Studies with high imputation RMSE (S2, S6, S9) and missing rate 61-70% also have low confidence. **Figure 2** shows that training with joined dataset stabilizes and improves imputation RMSE. Prediction results depend more on predictors than imputation methods.
- **Domain adaptation - withdrawn patients:** We test on withdrawn patients and record in **Figure 3** RMSE for each combination of imputation and prediction methods used. Prediction RMSE is worse than on test dataset. All predictors trained on single datasets give meaningless predictions with RMSE higher than 4 points. Predictors achieve a 50% RMSE decrease when using the joined dataset.

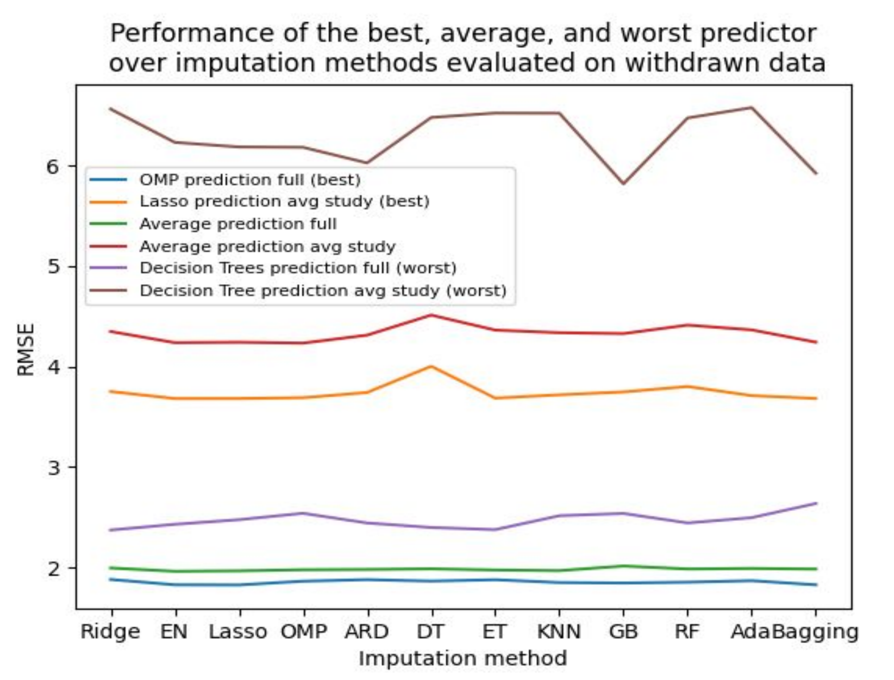
	Regression alg.	Full RMSE	Avg study RMSE
Baseline methods	Mean	5.758 ± 0.188	5.683 ± 3.916
	Median	5.832 ± 0.210	5.965 ± 4.275
	KNN	4.665 ± 0.202	4.837 ± 3.183
	Soft Impute	4.990 ± 0.188	5.298 ± 2.834
	SVD	6.527 ± 0.377	9.104 ± 5.293
	EM	7.960 ± 0.265	7.857 ± 5.626
	II DT	5.633 ± 0.085	6.131 ± 3.489
	II Extra Trees	3.984 ± 0.177	4.300 ± 2.038
Proposed methods	II Linear	8.619 ± 1.947	Not valid
	II BR	4.056 ± 0.115	4.451 ± 2.830
	II Ridge	4.284 ± 0.398	5.882 ± 3.251
	II Elastic Net	3.886 ± 0.145	4.430 ± 2.992
	II Lasso	3.935 ± 0.129	4.410 ± 2.909
	II OMP	3.982 ± 0.229	4.291 ± 2.669
	II ARD	7.225 ± 3.549	8.863 ± 9.864
	II KNN	4.206 ± 0.188	4.716 ± 2.829
	II RF	4.059 ± 0.204	4.213 ± 2.183
	II GB	3.913 ± 0.148	4.457 ± 2.379
	II Ada Boost	3.903 ± 0.180	4.476 ± 3.322
	II Bagging	$\mathbf{3.809 \pm 0.181}$	$\mathbf{4.195 \pm 2.869}$

TABLE I
IMPUTATION EVALUATED ON JOINED DATASET (FULL RMSE) AND AVERAGE RMSE OF IMPUTATION ON SEPARATE STUDIES (AVG STUDY RMSE)

Study	Imp. RMSE	Pred. RMSE	Coverage
S1	2.885 ± 0.842	1.029 ± 0.311	25%|1.7%
S2	3.983 ± 0.201	0.9034 ± 0.041	28%|4.1%
S3	3.229 ± 0.164	1.335 ± 0.192	60%|16%
S4	3.018 ± 0.087	1.359 ± 0.062	51%|39%
S5	2.805 ± 0.188	1.529 ± 0.030	24%|5.3%
S6	4.942 ± 1.207	1.439 ± 0.087	30%|2.1%
S7	1.594 ± 0.120	1.618 ± 0.358	35%|8.0%
S8	3.661 ± 0.549	1.814 ± 0.475	19%|2.1%
S9	11.639 ± 1.284	Impossible	23%|3.0%

TABLE II
PERFORMANCE OF IMPUTATION (BAGGING WITH ELASTICNET) AND PREDICTION (BAYESIAN REGRESSION) ON IMPUTED DATA FOR EACH STUDY.





References
[1] K. R. Merikangas, J. P. He, M. Burstein, S. A. Swanson, S. Avenevoli, L. Cui, J. Swendsen, "Lifetime prevalence of mental disorders in U.S. adolescents: Results from the National Comorbidity Survey Replication– Adolescent Supplement (NCS-A)," Journal of the American Academy of Child and Adolescent Psychiatry, 2010, vol. 49, pp. 980-989. doi:10.1016/j.jaac.2010.05.017.
[2] B. H. Bitsko, J. R. Holbrook, R. M. Ghandour, S. J. Blumberg, S. N. Visser, R. Perou, J. Walkup, "Epidemiology and impact of healthcare provider diagnosed anxiety and depression among US children," Journal of Developmental and Behavioral Pediatrics, Jun 2018, vol. 39, no. 5, pp. 395-403, doi: 10.1097/DBP.0000000000000571.