An adversarial learning method for missing data imputation based on feature decomposition

Penghao Gao, Tianhao Li, and Dengfeng Li*

University of Electronic Science and Technology of China, Sichuan, China
lidengfeng@uestc.edu.cn

Abstract. Missing data imputation is one of the problems frequently encountered in data analysis, and the quality of the data affects the outcome of the data mining. When the percentage of missing data is small, the records with missing values can be discarded or processed manually. In practice, a certain percentage of missing data can not be avoidable. Manual processing of missing data would be very inefficient, and discarding whole records with missing values would result in a large amount of information being lost, thus making systematic deviation between incomplete and complete observations. In this paper, we propose an Singular Value Decomposition Variational Auto-Encoder Generative Adversarial Networks (SVD-VAEGAN) method for filling in missing values of time-series features in Electronic Health Records (EHRs) for random missing data. In order to privide more efficient data into the generative model, missing data is initially imputed by SVD, and then uses the trained model for the missing data generation by using VAE as a generator of GAN. The results show that our model outperforms existing models in terms of MSE and performs more efficiently in the mortality risk prediction task.

Keywords: Data imputation · Deep generative model · Electronic health records.

1 Introduction

In real world scenarios, missing data is commonly found. Missing data may be caused by not performing operations that result in data not being available, or by accidental omissions during operations. Especially in the case of data related to medical tests on patients, there are few effective ways to ensure that various laboratory tests on patients are available in real time, resulting in missing values for some of the measurements. For the researcher, missing values may affect the accuracy of the model as well as its operational performance, or even lead to errors in building the model, which can lead a detrimental prediction and impact on productive life.

Existing data mining algorithms always need complete data. When the proportion of missing values is small, they can be filled in manually or simply discarded. However, when a large proportion of data is missing, the traditional approaches fill missing values with the mean or linear interpolation which makes

the data coarse and lack in individual differentiation. Coarse data can build the biased models, so finding suitable methods for missing data imputation has been a pressing problem.

Deep generative models (DGM) in recent years have mainly focused on the indeed filling of graphs [15, 29], while the missing filling of structured data imputed by DGM needs further research [28]. This paper focuses on designing a missing value filling method based on generative adversarial networks named SVD-VAEGAN (Singular Value Decomposition Variational Auto-Encoder Generative Adversarial Networks).

The main structure of the model as Fig.1, the algorithm firstly employs SVD [26] to perform a initial imputation of the missing data X named Y, and then goes through the Generative Adversarial Networks (GAN) [12] with the generator constituted by VAE [14]. The generative model and the discriminative model are continuously confronted through the GAN in the training process until convergence, and the generative model with convergence parameters is used to fill the missing values named M.

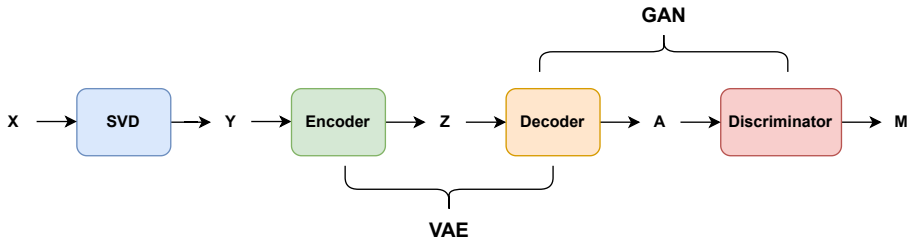

Fig. 1. The structure of proposed SVD-VAEGAN model.

Here are some contributions as follows:

- Designing SVD-VAEGAN for structured medical missing data in Electronic Health Records (EHRs), our model outperforms the traditional model on MSE in imputation task.
- In order to explore the performance of the model under different missing proportions, our model outperforms existing models in terms of MSE in most different proportions.
- After using different methods to impute the data, we put imputed data into different classifiers to discuss the mortality prediction. The results show our model achieves higher evaluation scores than traditional methods on the mortality prediction task.

2 Related works

Missing data is a complex problem in many research areas. However, in practical scenarios such as economic data and EHRs, missing data is common and the

proportion of missing data is often high. Manual processing of missing data can be inefficient, and trashing leads to the loss of a large amount of information and causes biased analysis results.

2.1 Summary of data imputation

Missing data are classified into missing completely at random (MCAR), missing at random (MAR) and missing not at random (MNAR) [17], depending on the distribution of the missing data.

Most of the current methods for filling missing data are based on MCAR and MAR. The traditional techniques are some statistical methods such as mean filling, linear interpolation, etc. The mean filling method replaces the missing values with the mean of the observed variable values, but this method ignores the correlation of the data [22], and has an excessive filling error [3], and the linear interpolation method, although it considers the correlation of the data, can only be used for data with linear relationships [4]. Multivariate Imputation by Chained Equations (MICE) [27] fills the missing data by using iterative regression model. And for data with non-linearity, the Maximum likelihood method of great likelihood estimation can be filled [8], but the prior distribution of the variables needs to be considered. In recent years, with the development of artificial intelligence, machine learning methods have been used to deal with missing data in traffic, weather and medical fields with good results. Machine learning methods can fit non-linear relationships in missing data very well, which are beyond the reach of traditional methods.

As for deep learning, it focuses more on distributing more complex data. For example, MLP for filling missing data shows the superiority than traditional algorithms [24]. For images imputation methods [1, 2, 7, 15, 16], their algorithms play good filling effects for different missing parts of images and different object states. As for multivariate time series data, there are some effective models [10, 18, 23] imputating for them.

2.2 Deep generative models for data imputation

Generative models are a representative method of deep learning. Generative models can learn the underlying distribution of the sample data then generating data, which are the ideal model for imputing missing data with a better performance [21]. Representative of generative models is GAN [12] and its variants such as AmbientGAN [2] and MIWAE [19], learns the complex distributions of data and are often used for image imputation. As for structured data, GAMIN [29] and MIDA [11] achieve effective enhancement for both highly missing structured data and image data imputation, and GAIN [28] proposes a new framework that can successfully interpolate using incomplete data and achieve better results than any other algorithms.

2.3 Data imputation in medical domain

Medical data imputation, which usually includes medical images and structured data. Image imputation is challenging and necessary for medical data, because the missing pixels filling in image has a limit of the upper and lower grey scale. Structured data do not have obvious upper and lower limits and individual differences are obvious [9]. Traditional methods such as multiple interpolation, do not capture the structure of the temporal data in large-scale medical data [25], thus reducing the performance of the model. An adversarial learning method [13] suggests its superoirty, but there are alse few researches about deep learning for data imputation in medical domain.

3 Proposed method

In this section, we present the design of the model for MCAR, including the initial filling of the data using the SVD algorithm in the first stage and further filling using the adversarial algorithm of GAN in the second stage.

3.1 Definition of masking

The input data comes from complete time-series data with n feature dimensions and m series. The complete data is simulated with random missingness and the missing values are denoted by X. The computer does not recognise the null values, so the missing data is replaced with a 0 to produce the Init Matrix. The non-missing data is then replaced with a 1 to produce the mask matrix. Such a masking process in Fig.2 takes care of pre-processing the missing values while retaining the location information.

Fig. 2. The process of generating Mask Matrix.

3.2 Singular value decomposition

SVD process is used for the initial completion of the init matrix in the first stage. The idea of SVD is to decompose the matrix using the singular values and replaces the missing values with the product of the decomposed matrix. Since the SVD does not require the decomposed matrix to be square, assume that the init matrix is an matrix with $M \times N$. So the definition of matrix A by SVD as:

$$A = U\Sigma V^T, \tag{1}$$

where U is an $m \times m$ matrix, Σ is an $m \times n$ matrix with all elements zero except the diagonal. And each element on the diagonal is called a singular value, and V is an $n \times n$ matrix. U, Σ, V can be computed by eigendecomposition via $A \times A^T$, and Fig.3 shows the process.

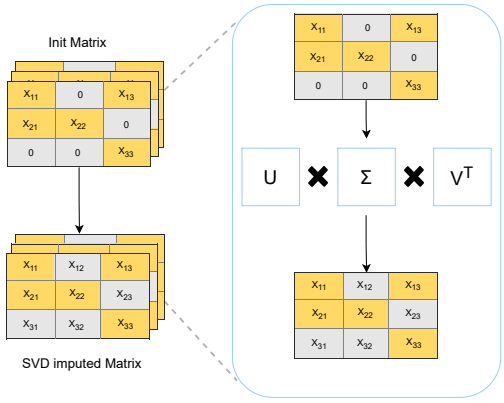

Fig. 3. The process of first stage imputation by SVD.

3.3 Generative Adversarial Nets

GAN consists of two main components, the generator and the discriminator. The generator is responsible for generating fake data, while the discriminator needs to discriminate between true and false data and train using the maximum-minimum game until equilibrium. The loss function of GAN is defined as follows:

$$\min_G \max_D V(G, D) = \mathrm{E}_{x \sim p_{data}(x)} \log D(x) + \mathrm{E}_{x \sim p_z(z)} \left[\log\left(1 - D\left(G\left(z\right)\right)\right)\right], \tag{2}$$

where G represents generator and D donates discriminator. x denotes the sample from data and z represents the latent varible. Our design of GAN employs VAE for the generator and Long Short-Term Memory (LSTM) for the discriminator to discriminate the temporal data.

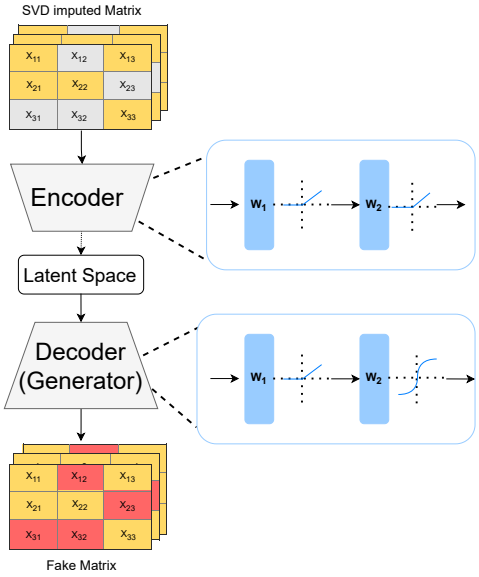

Fig. 4. The process of generating data by VAE-GAN.

3.4 Variational Autoencoder

In the GAN part, the generator is designed as a VAE. VAE is able to learn the parametric distribution in the latent space compared to autoencoder (AE), which ensures that the latent space is regularized and has a more powerful generative power than AE [14]. VAE is made up of an encoder model $Q_\phi(z|x)$ generates latent vector z from input x and a decoder $P_\theta(x|z)$ for reconstructing the input from the latent code z. The loss function of VAE consists of two parts, loss reconstruction and regularization, which is shown in following equation:

$$\mathrm{L}(\theta, \phi) = \mathrm{E}_{z\sim Q}\log P_\theta(x|z) - D_{KL}(Q_\phi(z|x)||P_\theta(z)), \tag{3}$$

where D_{KL} estimates the distance between these two densities by KL divergence.

4 Experiments

In this section we first present the sources of the data and the corresponding processing methods, then we describe the evaluation matrix for the different tasks and the algorithms for comparison, and finally we introduce the design of the experimental parameters.

4.1 Data extraction and processing

The dataset used in this paper is MIMIC-IV, a dataset containing real patient EHRs from tertiary academic medical centres in Boston, USA.

We initially use SQL to extract the basic information and laboratory test values of the patients, with a total of 70 features. Each patient is also divided into six time points, according to the time of admission and discharge. A preliminary sample population is eventually obtained.

Due to the large number of missing values, we remove variables with more than 50% missing values, and patients with more than 50% missing measures. We ultimately leave 9926 patients and nine variables, containing Weight, Heart Rate, SBP (Systolic BP), MBP (Mean BP), DBP (Diastolic BP), Temperature, Spo2 (O2 saturation pulseoxymetry), Gcs (Glasgow Coma Scale), Urine. The summary of the variables is shown in following Table 1.

Table 1. Summary of experimental variables.

Quantile	Min	Max	Mean	Std.
Weight	5.70	220.00	81.63	45.43
Heart rate	27.75	177.11	84.55	16.93
SBP	39.50	228.00	118.87	18.04
MBP	14.00	202.25	78.25	12.54
DBP	14.75	161.67	62.85	12.43
Temperature	29.90	41.33	36.92	0.60
Spo2	90.00	100.00	96.86	2.45
Gcs	3.00	15.00	14.52	1.39
Urine	0.00	3000.00	350.00	304.62

From Table 1, we can know that there is a big difference between the maximum and minimum values of urine. Temperature and Spo2 are relatively stable values. Although the distribution of the values of the different variables poses a challenge for imputation, GAN can still learn their distributions.

4.2 Feature processing

We fill the empty value with the mean of the measurements to obtain the complete data. We code the gender with one-hot. The rest of the variables are normalized using maximum-minimum as follows:

$$X'_{i,j} = \frac{X_{i,j} - X_{i,j_{min}}}{X_{i,j_{max}} - X_{i,j_{min}}} \tag{4}$$

where $X'_{i,j}$ represents the normalized value, i denotes the number of the patient, and j denotes the variable.

4.3 Evaluation matrix

In the task of predicting missing data, we use mean squared error (MSE) as our evaluation matrix. MSE can effectively portray the error between the predicted

and true values and calculated as follows:

$$\mathrm{MSE} = \sum_{n=1}^{k} (y_n - \hat{y}_n)^2, \tag{5}$$

where y_n represents the true value, and $\hat{y}_n$ represents the predicting value.

In the mortality risk prediction task, we chose indicators such as Accuracy, Precision and F_1 as evaluation indicators. Accuracy describes the proportion of samples with correct predictions to the overall sample. Recall is also called sensitivity in conception, which is concerned with the proportion of positive samples being predicted correctly, and the formula is as follows:

$$\mathrm{Recall} = \frac{TP}{TP + FN} \tag{6}$$

where $TP + FN$ represents the number of positive examples in the sample,and TP represents the number of positive examples that are predicted to be correct. The F_1 Score, also known as the balanced F-score, is defined as the summed average of Precision and Recall. The value of F_1 ranges from 0 to 1, and the value closer to 1 means the classifier performs better. The calculation of F_1 is as follows:

$$\mathrm{F}_1 = \frac{2 \times \mathrm{Precision} \times \mathrm{Recall}}{\mathrm{Precision} + \mathrm{Recall}}. \tag{7}$$

4.4 Comparison methods

To compare the effectiveness of our proposed algorithm, we compare other imputation methods, including following methods: (1) mean interpolation, which uses the mean of data to impute the missing value, (2) linear interpolation, which imputes the missing data depending on multivariate linear relationship, (3) matrix completion, [20] which replaces the missing elements with those obtained from a soft-thresholded SVD, (4) MiceForest, which is a method to deal with missing values based on repeated simulation and Monte Carlo method, (5) MissForest [5], which is an iterative imputation method based on a random forest.

4.5 Hyperparameter

To accurately evaluate the effectiveness of our proposed model, we divide the data into training and test sets according to the ratio of 8:2 under different random number seeds for repeat experiments.

In the VAE-generator nets, we use hidden layers with 54,27,9 neurons respectively, and the discriminator nets has 27 hidden neurons. The activation functions are ReLu in generator and discriminator, and Softmax is employed in output layers of discriminator. As for optimizer, we select Adam with a 0.003 learning rate.

5 Results

To verify the robustness of the proposed model, each of our experiments is repeated ten times to take the mean for reporting. The results of imputation are shown in Table 2. It shows the imputation performance in terms of MSE under different missing rates.

As we can see from Table 2, our model S-VAE-GAN significantly outperforms each benchmark in every missing rate. Matrix and Miceforest algorithms do not perform well on MSE, they are an order of magnitude higher compared to other algorithms. All algorithms except Matrix achieved their best performance at 30% missing rate, with our model attaining an MSE of 0.0145 at its best. Mean and Linear methods perform similarly at low missing rates, and Mean performs better than Linear as the missing rate increased. At the same time, if only the SVD algorithm of the first stage is used, the MSE will be 10 times larger than the final algorithm. The results show the effectiveness of our algorithm design.

Table 2. Imputation performance under different missing rates.

Missing rate	5%	10%	20%
Matrix	0.1824(0.0024)	0.1806(0.0054)	0.1520(0.0037)
Miceforest	0.1119(0.0079)	0.1042(0.0071)	0.1096(0.0035)
Missforest	0.0209(0.0038)	0.0176(0.0017)	0.0156(0.009)
Mean	0.0204(0.0000)	0.0184(0.0000)	0.0156(0.0000)
Linear	0.0214(0.0047)	0.0176(0.0018)	0.0154(0.0014)
S-VAE-GAN	**0.0195(0.0031)**	**0.0175(0.0015)**	**0.0151(0.0009)**
Missing rate	30%	40%	50%
Matrix	0.1313(0.0033)	0.1132(0.0015)	0.0961(0.0012)
Miceforest	0.1113(0.0024)	0.1154(0.0017)	0.1149(0.0008)
Missforest	0.0149(0.0008)	0.0172(0.0004)	0.0178(0.0004)
Mean	0.0150(0.0000)	0.0163(0.0000)	0.0168(0.0000)
Linear	0.0173(0.0007)	0.0224(0.0005)	0.0270(0.0005)
S-VAE-GAN	**0.0145(0.0006)**	**0.0158(0.0003)**	**0.0163(0.0004)**

Complemented data is obtained and can be used for a variety of tasks, such as mortality risk prediction using complemented medical data.

We use nine variables with 20% missing values to make complements using the S-VAE-GAN and bench algorithms, and the completed data is used to make predictions about the risk of death for patients by using Decision Tree (DT) and XGBoost [6], and the predictions are shown in Table 3 and Table 4. Data can be used in a variety of tasks after imputation. For example, medical data can be used in classification tasks. To test the validity of the imputed data applied to the classification task, we use the data imputed by S-VAE-GAN and other benchmarks into the medical classification task and test the accuracy. The missing rate is 20% and the results are shown in Table 3 and Table 4.

As we can see from Table 3 and Table 4, our model achieves optimal performance on all three evaluation metrics. From F_1 and recall, the linear in-

Table 3. Predicting performance under DT.

Algorithm	Accuracy	F1 score	Recall
Matrix	0.8277(0.0177)	0.4833(0.0719)	0.4833(0.0675)
Miceforest	0.8222(0.0245)	0.4522(0.0857)	0.5306(0.1035)
Missforest	0.8167(0.0088)	0.4407(0.0437)	0.4333(0.0434)
Mean	0.8111(0.0173)	0.4102(0.0370)	0.4363(0.0412)
Linear	0.8056(0.0124)	0.5000(0.0656)	0.5333(0.0649)
S-VAE-GAN	**0.8639(0.0069)**	**0.5763(0.0479)**	**0.5665(0.0580)**

Table 4. Predicting performance under XGBoost.

Algorithm	Accuracy	F1 score	Recall
Matrix	0.8500(0.0103)	0.5045(0.0578)	0.5873(0.0909)
Miceforest	0.7639(0.0113)	0.2456(0.0284)	0.2830(0.0373)
Missforest	0.7583(0.0193)	0.3063(0.0590)	0.3091(0.0508)
Mean	0.7500(0.0242)	0.2881(0.0460)	0.3035(0.0508)
Linear	0.8027(0.0105)	0.3469(0.0496)	0.4219(0.0441)
S-VAE-GAN	**0.8694(0.0159)**	**0.5155(0.0190)**	**0.5434(0.0379)**

terpolation results also outperform the other models besides our model. The performance of SVD is not as good as our model's, only improving on Recall of XGBoost. As the number of patient deaths is much greater than the number of survivors, the differences introduced by the unbalanced sample are more pronounced.

6 Conlcusion

We propose a two-step model for temporal missing data imputation named SVD-VAEGAN. The missing values are initially complemented by SVD, then the VAE-GAN is used to further learn the sample distribution to achieve a better performance. Our model is verified in time-series EHRs and get the lower MSE than the bench model in imputation task. We further use the completed data to get better performance than bench model in death analysis and prediction task.

Acknowledgements This work was supported by the National Key Research and Development Program of MOST of China under Grant 2018AAA0101003, and National Natural Science Foundation of China (Grant No. 71901050).

Bibliography

[1] Becker, P., Pandya, H., Gebhardt, G.H.W., Zhao, C., Taylor, C.J., Neumann, G.: Recurrent kalman networks: Factorized inference in high-dimensional deep feature spaces. In: Chaudhuri, K., Salakhutdinov, R. (eds.) Proceedings of the 36th International Conference on Machine Learning, ICML 2019, 9-15 June 2019, Long Beach, California, USA. Proceedings of Machine Learning Research, vol. 97, pp. 544–552. PMLR (2019)

[2] Bora, A., Price, E., Dimakis, A.G.: Ambientgan: Generative models from lossy measurements. In: 6th International Conference on Learning Representations, ICLR 2018, Vancouver, BC, Canada, April 30 - May 3, 2018, Conference Track Proceedings. OpenReview.net (2018)

[3] Brown, C.H.: Asymptotic comparison of missing data procedures for estimating factor loadings. Psychometrika **48**(2), 269–291 (1983)

[4] Buck, S.F.: A method of estimation of missing values in multivariate data suitable for use with an electronic computer. Journal of the Royal Statistical Society: Series B (Methodological) **22** (1960)

[5] Bühlmann, P.: Missforest—non-parametric missing value imputation for mixed-type data. Bioinformatics **28**(1), 112–118 (2012)

[6] Chen, T., Guestrin, C.: Xgboost: A scalable tree boosting system. In: the 22nd ACM SIGKDD International Conference (2016)

[7] Dalca, A.V., Bouman, K.L., Freeman, W.T., Rost, N.S., Golland, P.: Medical image imputation from image collections. IEEE Transactions on Medical Imaging **PP**(99), 1–1 (2018)

[8] Dempster, A.P., Laird, N.M., Rubin, D.B.: Maximum likelihood from incomplete data via the em algorithm. Proceedings of the Royal Statistical Society **39**(1), 1–22 (1977)

[9] Donald, B., Rubin, Nathaniel, Schenker: Multiple imputation in health-are databases: An overview and some applications. Statistics in Medicine **10**(4), 585–598 (1991)

[10] Fortuin, V., Baranchuk, D., Rätsch, G., Mandt, S.: GP-VAE: deep probabilistic time series imputation. In: Chiappa, S., Calandra, R. (eds.) The 23rd International Conference on Artificial Intelligence and Statistics, AISTATS 2020, 26-28 August 2020, Online [Palermo, Sicily, Italy]. Proceedings of Machine Learning Research, vol. 108, pp. 1651–1661. PMLR (2020)

[11] Gondara, L., Ke, W.: Mida: Multiple imputation using denoising autoencoders. Springer, Cham (2018)

[12] Goodfellow, I.J., Pouget-Abadie, J., Mirza, M., Xu, B., Warde-Farley, D., Ozair, S., Courville, A., Bengio, Y.: Generative Adversarial Networks. arXiv e-prints arXiv:1406.2661 (Jun 2014)

[13] Hallaji, E., Razavi-Far, R., Palade, V., Saif, M.: Adversarial learning on incomplete and imbalanced medical data for robust survival prediction of liver transplant patients. IEEE Access **9**, 73641–73650 (2021)

[14] Kingma, D.P., Welling, M.: Auto-encoding variational bayes. arXiv.org (2014)
[15] Lee, D., Kim, J., Moon, W., Ye, J.C.: Collagan: Collaborative GAN for missing image data imputation. In: IEEE Conference on Computer Vision and Pattern Recognition, CVPR 2019, Long Beach, CA, USA, June 16-20, 2019. pp. 2487–2496. Computer Vision Foundation / IEEE (2019). https://doi.org/10.1109/CVPR.2019.00259
[16] Li, S.C., Jiang, B., Marlin, B.M.: Misgan: Learning from incomplete data with generative adversarial networks. In: 7th International Conference on Learning Representations, ICLR 2019, New Orleans, LA, USA, May 6-9, 2019. OpenReview.net (2019)
[17] Little, R., Rubin, D.B.: Statistical Analysis with Missing Data, Second Edition. Statistical Analysis with Missing Data (2002)
[18] Luo, Y., Zhang, Y., Cai, X., Yuan, X.: Egan: End-to-end generative adversarial network for multivariate time series imputation. In: Twenty-Eighth International Joint Conference on Artificial Intelligence IJCAI-19 (2019)
[19] Mattei, P., Frellsen, J.: MIWAE: deep generative modelling and imputation of incomplete data sets. In: Chaudhuri, K., Salakhutdinov, R. (eds.) Proceedings of the 36th International Conference on Machine Learning, ICML 2019, 9-15 June 2019, Long Beach, California, USA. Proceedings of Machine Learning Research, vol. 97, pp. 4413–4423. PMLR (2019)
[20] Mazumder, Rahul, Hastie, Trevor, Tibshirani, Robert: Spectral regularization algorithms for learning large incomplete matrices. Journal of Machine Learning Research (2010)
[21] Mccoy, J.T., Kroon, S., Auret, L.: Variational autoencoders for missing data imputation with application to a simulated milling circuit. IFAC-PapersOnLine **51**(21), 141–146 (2018)
[22] Mortaza, Jamshidian, Peter, M., Bentler: Ml estimation of mean and covariance structures with missing data using complete data routines. Journal of Educational and Behavioral Statistics (2016)
[23] Shan, S., Oliva, J.B.: NRTSI: non-recurrent time series imputation for irregularly-sampled data. CoRR **abs/2102.03340** (2021), https://arxiv.org/abs/2102.03340
[24] Silva-Ramírez, E.L., Pino-Mejías, R., López-Coello, M., de-la Vega, M.D.C.: Missing value imputation on missing completely at random data using multilayer perceptrons. Neural Networks the Official Journal of the International Neural Network Society **24**(1), 121–129 (2011)
[25] Sterne, J., White, I.R., Carlin, J.B., Spratt, M., Carpenter, J.R.: Multiple imputation for missing data in epidemiological and clinical research: potential and pitfalls. BMJ (online) **338**(jun29 1), b2393 (2009)
[26] Troyanskaya, O.G., Cantor, M.N., Sherlock, G., Brown, P.O., Hastie, T., Tibshirani, R., Botstein, D., Altman, R.B.: Missing value estimation methods for DNA microarrays. Bioinform. **17**(6), 520–525 (2001). https://doi.org/10.1093/bioinformatics/17.6.520
[27] White, I.R., Royston, P., Wood, A.M.: Multiple imputation using chained equations: Issues and guidance for practice. Statistics in Medicine **30**(4) (2011)

[28] Yoon, J., Jordon, J., van der Schaar, M.: GAIN: missing data imputation using generative adversarial nets. In: Dy, J.G., Krause, A. (eds.) Proceedings of the 35th International Conference on Machine Learning, ICML 2018, Stockholmsmässan, Stockholm, Sweden, July 10-15, 2018. Proceedings of Machine Learning Research, vol. 80, pp. 5675–5684. PMLR (2018)
[29] Yoon, S., Sull, S.: Gamin: Generative adversarial multiple imputation network for highly missing data. In: 2020 IEEE/CVF Conference on Computer Vision and Pattern Recognition (CVPR) (2020)