Unsupervised identification of atypical medication orders: A GANomaly-based approach

Maxime Thibault, B.Pharm, M.Sc[1]; Pierre Snell, M.Sc[2]; Audrey Durand, PhD[2,3]
[1] CHU Sainte-Justine, [2]Université Laval, [3]Mila – Quebec AI Institute

Background.

The analysis of medication orders using machine learning has either focused on single orders [1-3], or on the prediction of future orders for a patient [4-6]. To our knowledge, no study has attempted to detect anomalies in lists of currently active medications for a patient (pharmacological profiles). For this purpose, we propose an adaptation of GANomaly [7], which is an adversarially-trained autoencoder. The general idea is to learn a latent space in such way that anomalies can be identified by computing the loss between latent representations from two encoders respectively taking as input the original data and its reconstruction (encoder loss). The practical use of this model would be to identify profiles most likely to contain anomalies and triage them for pharmacist review, in order to reduce the cognitive burden of the medication order review process. In its current state, this process involves pharmacists reviewing a majority of profiles containing no errors or problems [8-13]. The objectives of this study were to determine the feasibility of this technique.

Methods.

The dataset used for this study consists of all medication orders placed between 2005 and 2018 inclusively, contained in the pharmacy database of CHU Sainte-Justine, a tertiary-care mother-and-child academic hospital center located in Montreal, Canada. Access to the data was authorized in conformity with local requirements. We preprocessed the data to reconstruct pharmacological profiles, defined as sets of active medications for a patient at any time point where no order was placed in the following hour. These were represented as multi-hot vectors. The dataset does not include labels as to the atypical nature of orders or profiles. For this reason, we selected an unsupervised autoencoder-based GAN technique focused on anomaly detection, namely GANomaly. Given the binary nature of the data, we adapted GANomaly by replacing the L1 loss by a binary crossentropy contextual loss. The original technique assumes that the dataset would include only samples without anomalies. However, we hypothesized that the large proportion of orders following common patterns would allow a properly calibrated model to yield higher encoder losses for atypical profiles.

Data from years 2005 to 2017 (inclusively), that is 989,766 sets containing (mean ± std) 9.43 ± 6.88 drugs derived from 236,429 patient encounters including 4329 different drugs, was used for training and validation, leaving 73,407 sets of 10.11 ± 7.56 drugs from 2018 for testing. As a starting point, we seeked to establish a base autoencoder architecture that demonstrated a capacity to reconstruct the multi-hot vectors with more than 80% accuracy when training on a single year (2014, 2015, or 2016) and validating on the next year. Since published literature suggests that there may be data drift through time on this type of data [14], we verified this effect with 3-fold cross-validation, using years 2015, 2016 and 2017 as validation splits and training on the preceding 1 to 10 years. Based on results from an in-house preliminary study where pharmacists flagged about 10% of orders as atypical, we then tuned the autoencoder architecture until we reached a proportion of orders predicted as atypical (i.e. appearing in the original input but absent from the reconstruction) close to 10%. The adversarial training hyperparameters were tuned following the same cross-validation scheme. We finally performed an adversarial training of the resulting autoencoder model and evaluated it on the test set.

Results.

The optimal autoencoder architecture was composed of 3 dense layers (size 256, 64 and 256), with SELU activation and a dropout ratio of 0.1. This structure yielded a (mean ± std) validation reconstruction accuracy of 88.8 ± 1.1%. Increasing the training data volume up to 10 years consistently improved performance. The optimal discriminator structure was 2 dense layers (size 128 and 64) with ReLU activation, batch normalization, and a dropout ratio of 0.1. Adversarial training was performed using the Adam optimizer with a learning rate of 10^{-6} for the discriminator and 10^{-3} for the adversarial model, and a batch size of 256. The differences between the encoder and discriminator and the reduced discriminator learning rate helped prevent mode collapse. Loss weights were 100 for contextual loss, 1 for adversarial loss, and 2 for encoder loss. The final model was trained for 21 epochs based on early stopping from validation loss. On the test set, this yielded a reconstruction accuracy of 0.87, an area under precision-recall of 0.88, and 11% of orders predicted as atypical. The lowest median encoder loss per profile was in departments which follow the most predictable patterns with a large proportion of orders being protocolized (e.g. nursery, obstetrics, NICU). Conversely, higher encoder losses were observed in departments which more frequently treat patients with complex

pharmacological profiles (e.g. PICU, long term care, oncology). A pharmacist reviewed 50 samples per department from the test set, which qualitatively showed that pharmacological profiles with a low encoder loss corresponded well to routine orders and protocols.

Conclusion.

At this point only unsupervised training and evaluation was performed. The adapted GANomaly is currently being evaluated in a prospective study comparing pharmacists' evaluation of medication profiles with model predictions.

References

1. Woods AD, Mulherin DP, Flynn AJ, *et al.* Clinical decision support for atypical orders: detection and warning of atypical medication orders submitted to a computerized provider order entry system. *J Am Med Inform Assoc* 2014; **21**: 569–73. doi:10.1136/amiajnl-2013-002008
2. Flynn AJ, Milstein JA, Boisvert P, *et al.* ScriptNumerate: A Data-to-Advice Pipeline using Compound Digital Objects to Increase the Interoperability of Computable Biomedical Knowledge. *AMIA Annu Symp Proc* 2018;**2018**:440–9.
3. Santos HDP dos, Ulbrich AHDPS, Woloszyn V, *et al.* DDC-Outlier: Preventing Medication Errors Using Unsupervised Learning. *IEEE J Biomed Health Inform* 2019; **23**: 874–81. doi:10.1109/JBHI.2018.2828028
4. Thibault M, Lebel D. An application of machine learning to assist medication order review by pharmacists in a health care center. *medRxiv* 2019;:19013029. doi:10.1101/19013029
5. Chen JH, Goldstein MK, Asch SM, *et al.* Predicting inpatient clinical order patterns with probabilistic topic models vs conventional order sets. *J Am Med Inform Assoc* 2017; **24**: 472–80. doi:10.1093/jamia/ocw136
6. Hunter-Zinck HS, Peck JS, Strout TD, *et al.* Predicting emergency department orders with multilabel machine learning techniques and simulating effects on length of stay. *J Am Med Inform Assoc* 2019; **26**: 1427–36. doi:10.1093/jamia/ocz171
7. Akcay S, Atapour-Abarghouei A, Breckon TP. GANomaly: Semi-Supervised Anomaly Detection via Adversarial Training. Published Online First: 17 May 2018.https://arxiv.org/abs/1805.06725v3
8. Mahoney CD, Berard-Collins CM, Coleman R, *et al.* Effects of an integrated clinical information system on medication safety in a multi-hospital setting. *Am J Health Syst Pharm* 2007; **64**: 1969–77. doi:10.2146/ajhp060617
9. Flynn AJ. Opportunity cost of pharmacists' nearly universal prospective order review. *Am J Health Syst Pharm* 2009; **66**: 668–70. doi:10.2146/ajhp070671
10. Dakwa DS, Marshall VD, Chaffee BW. The impact of drug order complexity on prospective medication order review and verification time. *J Am Med Inform Assoc* 2020; **27**: 284-93. doi:10.1093/jamia/ocz188
11. Poikonen J. An informatics perspective on nearly universal prospective order review. *Am J Health Syst Pharm* 2009; **66**: 704–5. doi:10.2146/ajhp080410
12. Tribble DA. Automating order review is delegation, not abdication. *Am J Health Syst Pharm* 2009; **66**: 1078–9. doi:10.2146/ajhp090095
13. Gorbach C, Blanton L, Lukawski BA, *et al.* Frequency of and risk factors for medication errors by pharmacists during order verification in a tertiary care medical center. *Am J Health-Syst Pharm* 2015; **72**: 1471–4. doi:10.2146/ajhp140673
14. Chen JH, Alagappan M, Goldstein MK, *et al.* Decaying relevance of clinical data towards future decisions in data-driven inpatient clinical order sets. *Int J Med Inf* 2017;**102**:71–9. doi:10.1016/j.ijmedinf.2017.03.006