UTHealth

The University of Texas Health Science Center at Houston

School of Biomedical Informatics

Applications of Neural Ordinary Differential Equations in Deep Learning Models Based on Irregularly-sampled Time Series Electronic Health Record Data

Bingyu Mao

BMI 7303: Critical Review of Biomedical Informatics Literature Seminar

The School of Biomedical Informatics |The University of Texas Health Science Center at Houston

Please contact the first author via email: Bingyu.Mao@uth.tmc.edu

INTRODUCTION

Analyzing EHR (electronic health record) data allows us to improve clinical decision support and predict clinical processes for specific conditions. Nevertheless, time-series EHR data are always incomplete and irregularly sampled.

The topic of this paper is applications of neural ODE in deep learning models (e.g., recurrent neural networks or RNNs) based on irregularly-sampled time series EHR data. Figure 1 shows the main topics of the 16 articles reviewed including the baseline model neural ODE from Chen et al. (2018) and how they related to each other.

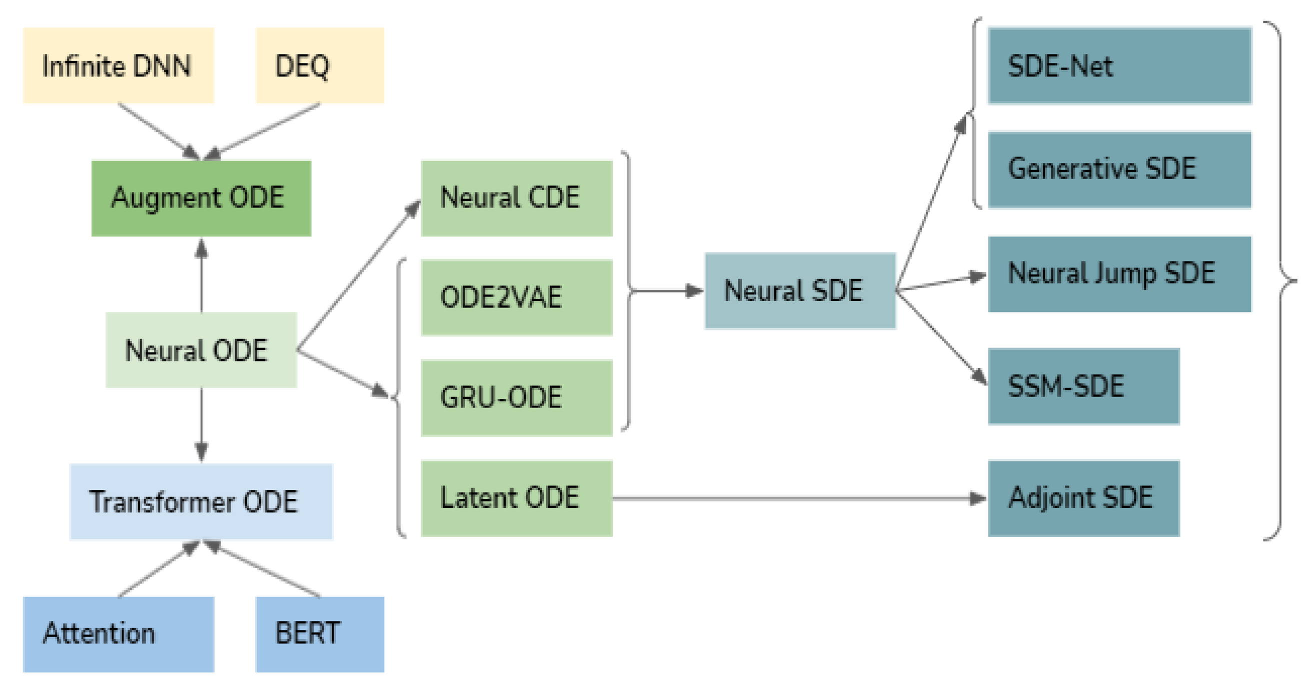

Figure 1: How ODE-based models related to the neural ODE model and also how they are related to each other

METHOD

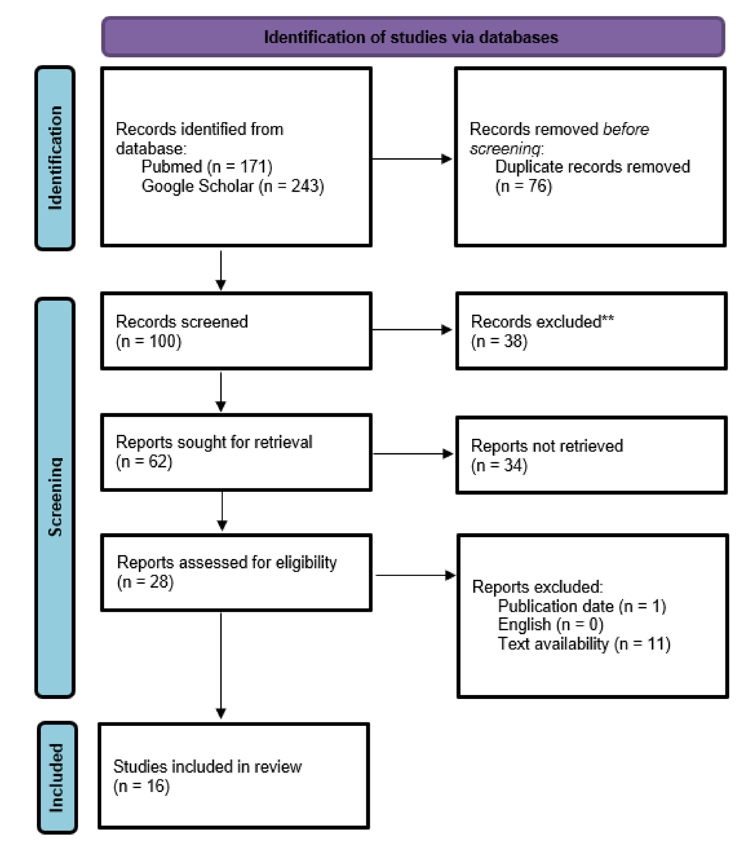

Figure 2: The PRISMA diagram

CONCLUSION

The current focus of my research is to develop deep learning models that take time-series data from EHRs and predict drug concentrations based on it, in which the concentration distribution of the drug itself conforms to an ODE.

Considering the specific data for my project, it is not necessary to utilize SDE-based methods. The latent ODE solves the problem from the nature of the ODE, it seems worthwhile to use it when the dataset is not too large. Alternatively, a neural CDE model can be employed, which smooths out the trajectory of the hidden state by using a natural cubic spline. As well, augmented ODEs may eventually replace traditional neural ODEs.

SUMMARY & DISCUSSION

Table 1: Summary of models

Type	Summary	Description	Future study
Neural ODE	Baseline model.	Updates hidden states continuously, but the model's trajectory cannot be adjusted based on observations.	Add uncertainty into the trajectory (ODE-related models); add random noise to fit real-world data (SDE-related models).
ODE-related	**Neural CDE**	Could smooth the invisible trajectory.	The latent ODE and the neural CDE have outperformed other models based on neural ODEs. It is essential to utilize latent ODEs on datasets that are not particularly large, despite their higher computational complexity and ability to lead to optimal hyperparameter determination. There are the same limitations of the neural CDE as that of the potential ODE: lower speed and more parameters.
	ODE2VAE (Variational autoencoder)	Is a latent second-order ODE model for high-dimensional sequential data.	
	GRU-ODE (Gated recurrent units)	Is the continuous-time version of the GRU.	
	Latent ODE	Can naturally handle arbitrary time gaps be- tween observations.	
SDE-related (Stochastic differential equation)	SDE-Net	A new method for quantifying uncertainties of DNNs (deep neural networks) from a dynamical system perspective.	Most SDE-related models are not "real" SDE, so the best model should be the adjoint SDE. When comparing the potential SDE model to the potential ODE model when using the same data set, Li et al. (2020) report that the test MSE (Mean square error) of the possible SDE model is slightly lower. It is difficult to determine under what circumstances neural SDE models will be more effective than neural ODE models because SDE models are more complex, memory, and time-intensive.
	Generative SDE	Uses numerical SDE solvers to generate samples.	
	Neural Jump SDE	Provides a data-driven approach to learn continuous and discrete dynamic behavior.	
	SSM-SDE (State space model)	A new model of insulin-glucose dynamics for forecasting blood glucose in type 1 diabetics.	
	Adjoint SDE	Allows time-efficient and constant-memory computation of gradients with high-order adaptive solvers.	
Other	Augment ODE	There exist functions that neural ODE cannot represent. Augmented ODEs are more expressive and stable models.	Eventually, augmented ODEs could replace traditional neural ODEs.
	Transformer ODE	Neural ODEs are typically defined using vector fields parameterized via Lipschitz networks. The authors (Kim et al., 2021) investigated the Lipschitz constant of self-attention.	As of yet, there is no literature linking the transformer-related model with the neural ODE network.

ACKNOWLEDGEMENTS

Thank Dr. Susan Fenton for guidance in BMI 7303. Thank Dr. Degui Zhi, Dr. Ziqian Xie, and Dr. Laila Irasmy Gindy Bekhet from Zhi Group for helpful discussions and suggestions.

REFERENCES

1. Brouwer, E.D., Simm, J., Arany, A., & Moreau, Y. (2019). GRU-ODE-Bayes: Continuous modeling of sporadically-observed time series. *NeurIPS*. https://arxiv.org/pdf/1905.12374.pdf
2. Chen, T. Q., Rubanova, Y., Bettencourt, J., and Duvenaud, D. K. (2018). Neural ordinary differential equations. *Advances in Neural Information Processing Systems 31*, pages 6571–6583. https://arxiv.org/pdf/1806.07366.pdf
3. Dupont, E., Doucet, A., & Teh, Y.W. (2019). Augmented Neural ODEs. NeurIPS. https://arxiv.org/pdf/1904.01681.pdf
4. Jia, J., & Benson, A.R. (2019). Neural Jump Stochastic Differential Equations. *NeurIPS*.
5. Kidger P., Morrill J., Foster J., and Lyons T. (2020). Neural controlled differential equations for irregular time series. *Advances in Neural Information Processing Systems*. https://arxiv.org/pdf/2005.08926.pdf
6. Kim, H., Papamakarios, G., & Mnih, A. (2021). The Lipschitz Constant of Self-Attention. ICML. https://arxiv.org/pdf/2006.04710.pdf
7. Kong, L., Sun, J., & Zhang, C. (2020). SDE-Net: Equipping Deep Neural Networks with Uncertainty Estimates. ArXiv, abs/2008.10546.
8. Li, X., Wong, T.L., Chen, R.T., & Duvenaud, D.K. (2020). Scalable Gradients for Stochastic Differential Equations. *ArXiv, abs/2001.01328*.
9. Miller, A.C., Foti, N.J., & Fox, E.B. (2020). Learning Insulin-Glucose Dynamics in the Wild. *ArXiv, abs/2008.02852*.
10. Rubanova, Y., Chen, T. Q., and Duvenaud, D. K. (2019). Latent ODEs for irregularly-sampled time series. *Advances in Neural Information Processing Systems*, pp. 5321–5331. https://arxiv.org/pdf/1907.03907.pdf
11. Song, Y., Sohl-Dickstein, J., Kingma, D.P., Kumar, A., Ermon, S., & Poole, B. (2021). Score-Based Generative Modeling through Stochastic Differential Equations. ArXiv, abs/2011.13456. https://arxiv.org/pdf/2011.13456.pdf
12. Yildiz, Ç., Heinonen, M., & Lähdesmäki, H. (2019). ODE2VAE: Deep generative second-order ODEs with Bayesian neural networks. NeurIPS. https://arxiv.org/pdf/1905.10994.pdf