

1. PLAN

1.1. **Catalyst (Intro).** Temporal networks pop up everywhere and would be really nice if we could model them.

1.2. **What methods are out there.** Exploration of whats there for temporal network modelling and what challenges they have faced

1.3. **Proposition for using DE (and problems that arise from this approach).**

1.4. **UDEs (we may not have much knowledge of the system so may use UDEs).** Exploration of its uses and limitations (Hopefully find some stuff about them being black boxes)

1.5. **Symbolic regression (Illustrate that it has been well researched and shown positive results).**

2. LITERATURE REVIEW

2.1. **Temporal Networks.** Temporal networks are present in many areas of research interest, ranging from symptom interactions for mental health[21, 19], to epidemiology[12], protein interactions[29, 5], and social networks[14, 7].

Generally attempts to model temporal networks either have the overall network structure static, but model the change in node state as in[19], or they model the changing structure of the network as in[11].

For example, [19] model the change in severity of symptoms over time, ie the state of the nodes, as a temporal network where the network structure remains constant. They use multilevel vector autoregression to model how a the current symptoms of people with paranoia might predict their future symptoms. These models were then used to create three networks that linked symptoms to eachother. Of particular

interest to this thesis is the temporal network created by having a fully connected network of all symptoms with self loops. The edge weights represent the extent to which the severity of each symptom at time (t) predicts the severity of itself and other symptoms at time $(t + 1)$. This framework keeps the overall structure of the network fixed throughout time and so is not particularly flexible, but is extremely human interpretable and can extrapolate, which would be very useful in many applications.

In [11], the authors look to model the change in network structure. They represent their temporal network as a series of interactions (i, j, t) , where i, j are nodes in the network and the interaction occurred at time t . The authors use a model to predict when a node will change from state 0 to state 1, notably, in their model the node never changes back; this is done to keep the model analytically tractable. To do this the model uses a sliding time window and if the fraction of interactions with nodes of state 1 within that time window exceeds a threshold, the node switches to state 1 as well. This model seems especially useful for understanding how ideas spread online for example, but by itself it cannot be used to predict the states of nodes in the future. To do this a model would also have to predict network structure.

[7] - TERGM (an expansion of ERGM models for temporal sequences) XXXX Ask Giulio

Pasino et al.[31] also model the change in structure of a temporal network. In this case the authors represent the temporal network as a series of adjacency matrices, each matrix representing an observation of the changing network. The authors look to model the structure of edges by employing time series techniques. The authors use a spectral embedding (SVD) to transform their temporal sequence into a latent position model, where nodes can be represented as points in a continuous, low dimensional space. With this series of embeddings, the authors employ a variety of time series techniques to predict the network structure at future time steps.

As we can see, temporal networks are relevant in a wide range of fields, and modelling has been the subject of ongoing and rich study. Generally modelling approaches fall into two broad categories modelling the evolving states of the nodes, or modelling the evolving structure of the network.

XXX more work needed, not happy with below but need to go through citations to fix it XXXX

temporal networks. One category are preposed with minimal context, offering as general an approach as possible[9, 11], and proposing more abstracted models which could be applied to almost any context. The other are models in which the context gave rise to the system by which it is modelled[24, 21, 17, 16].

2.2. UDEs and NODEs. In some networks the state of a node might influence the evolution of its neighbours. For example, in social networks we often observe the power law distribution of edges, where the more neighbours a node has, the more likely they are to gain more edges[10, 4]. We also see that the the state of symptoms (occurence, severity, and distress), can influence the states of other symptoms in patients undergoing chemotherapy [18, 32], and for mental health disorders [19]. As well as in ecological networks where the interactions of species are often represented as a network, and the populations of one species influence the population growth of another [2, 1].

With the observation that nodes in real world temporal networks influence others, differential equations or difference equations seem appropriate for modelling the evolution of nodes. One of the issues that the differential equation method encounters is the prevalence of discrete jumps of edges forming and decaying.

Because of this we may consider exploring the possibility of using difference equations. This approach may overcome the problem of discrete observations, by defining a progression function that will only give us discrete formations or decay of edges[7].

However, this imposes a significant limitation on the temporal granularity with which we can predict the state of the network. We would be limited to the granularity our data is collected at. This may be a problem if for example we had a series of weekly or monthly observations of networks we would not be able to predict interactions at each day. We can overcome both this limitation and the discrete jump limitation by using random dot product graphs.

Looking at the work of Passino et al.[31], the authors use random dot product graphs[15] to approximate a temporal network as a matrix of probabilities of an edge existing between two nodes. In contrast to edges, these probabilities can evolve continuously and so we can use differential equations to generate probabilistic networks at any temporal resolution we require.

Using the method in Passino et al.[31], we can treat the sequence as a dynamical system and then use differential equation modelling in a continuous space. With the aim of this framework to be as applicable to as many areas of research as possible, we note that there may not yet be enough knowledge in the field to perform traditional differential equation modelling. Because of this, we will use a recently developed tool a universal differential equation [22].

2.2.1. *Theory.* Developed by Rackauckas et al.[22], Universal differential equation (UDEs) are a novel neural network architecture that aims to take the best of both traditional neural network modelling and of flexible machine learning approaches to modelling. UDEs achieve this by combining some amount of domain knowledge in the form of a differential equation, with a neural network that is trained on the difference between the observed data and the prediction of the differential equation. Combining the two allows for the differential equation to be more flexible and for the neural network to only need to learn a theoretically simpler equation, as a part of the system has already been captured. This then allows the neural network to

be smaller, require less data, and be trained much faster than traditional neural network models. Given the flexibility of neural networks, the use of UDEs is natural when modelling a process that is not well understood[33].

Largely, the SciML ecosystem has been optimised for flexibility and efficiency with respect to models available and training performance. Given there does not seem to be any other ecosystem with this feature, it seemed like an obvious choice. What makes it good (Examples from [22])

2.2.2. Applications [Caution only glanced at these papers]. Whilst there has been much interest in implementing UDEs and neural ODEs, much of this work has been focussed on physics informed neural network and physics informed ordinary differential equations (PINNs and PINODEs) [26], [25], [28], [23], and on improving modelling of fluids [30] [34]. Along side this, and given UDEs have gained popularity in the last few years, there have been a number of studies exploring their usefulness in modeling the effect of restrictions due to COVID-19 on the virus' spread [20]. Although there has been a large amount of research into the usefulness of these types of models, to the best of this author's knowledge they have never been applied to the problem of predicting temporal social networks. This seems to be a natural fit for UDEs and neural ODEs, given we know relatively little about the processes that govern the evolution of online social networks, and the vast amounts of data that the providers of these services collect, data that can be used to train our neural networks.

Start by illustrating that there is a wide range of applications (Reasonably briefly for not relevant), but little research into UDEs aplicability for modeling how temporal networks behaveb

2.3. Symbolic Regression. Story is that there is a large search space of potential symbolic equations so most research has been looking to optimise this search.

The problem that symbolic regression aims to solve is to generate interpretable and "meaningful" equations from observed data. For example, both [6, 3] test their methods' capacity for obtaining a symbolic equation with physical simulation data, various pendula specifically. A "meaningful" equation in this case would be one that adheres to the laws of physics. To ensure this is the case, both use numerically calculated partial derivatives, a method that is common throughout the research.

This is relevant to all fields, and so its applications are incredibly varied. Various methods of symbolic regression have been used in manufacturing systems, chemical systems, and tumor research [8, 27, 13]. As well, research into improvements is still on going.

[33] also demonstrates that, with NODEs, two of the assumptions necessary for symbolic regression can be overcome. Those are the necessity for paired observations and derivatives. We are modelling a differential equation, and so will have access to the derivatives at any point they are defined at. As well as the assumption that the function can be expressed as a shallow tree of symbolic operations. XXX Unclear as to why this is satisfied XXX

Examples of methods include discovering equations governing the movement of simple harmonic and chaotic double pendula from data[6] (for the double pendulum searching took between 30-40 hours but when some knowledge of the system was added the time went to 7-8 hours). In this paper the authors present an evolutionary algorithm to generate multiple candidate functions that approximate the partial derivatives of each variable that are calculated directly from the data. A limitation of this is the fact that to have accurate derivatives, the data must be sampled very frequently. This seems integral to the accuracy of the found equation. Another approach to symbolic regression can be found in [3]. This paper proposes a method to find a symbolic equation that approximates a learned numerical equation rather than directly from data. The candidate equation is then tested against the numerical

equation in such a way that the difference between the two. Especially when using a neural network model as the numerical equation, it seems these process may not be useful if the initial conditions are outside of the scope of the training data.

As well numerous algorithms exist, with one of the most prominent being genetic programming XXXX research XXXX

REFERENCES

- [1] Vito Volterra. *Variazioni e fluttuazioni del numero d'individui in specie animali conviventi*. Vol. 2. Società anonima tipografica" Leonardo da Vinci", 1927.
- [2] Charles S Elton. *Animal ecology*. University of Chicago Press, 2001.
- [3] Josh Bongard and Hod Lipson. "Automated reverse engineering of nonlinear dynamical systems". In: *Proceedings of the National Academy of Sciences* 104.24 (2007), pp. 9943–9948.
- [4] Sanchit Garg, Trinabh Gupta, Niklas Carlsson, and Anirban Mahanti. "Evolution of an online social aggregation network: an empirical study". In: *Proceedings of the 9th ACM SIGCOMM Conference on Internet Measurement*. 2009, pp. 315–321.
- [5] Ruoming Jin, Scott McCallen, Chun-Chi Liu, Yang Xiang, Eivind Almaas, and Xianghong Jasmine Zhou. "Identifying dynamic network modules with temporal and spatial constraints". In: *Biocomputing 2009*. World Scientific, 2009, pp. 203–214.
- [6] Michael Schmidt and Hod Lipson. "Distilling free-form natural laws from experimental data". In: *science* 324.5923 (2009), pp. 81–85.
- [7] Steve Hanneke, Wenjie Fu, and Eric P Xing. "Discrete temporal models of social networks". In: (2010).

- [8] Birkan Can and Cathal Heavey. “Comparison of experimental designs for simulation-based symbolic regression of manufacturing systems”. In: *Computers & Industrial Engineering* 61.3 (2011), pp. 447–462.
- [9] Petter Holme and Jari Saramäki. “Temporal networks”. In: *Physics Reports* 519.3 (2012). Temporal Networks, pp. 97–125. ISSN: 0370-1573. DOI: <https://doi.org/10.1016/j.physrep.2012.03.001>. URL: <https://www.sciencedirect.com/science/article/pii/S0370157312000841>.
- [10] Xiaohan Zhao, Alessandra Sala, Christo Wilson, Xiao Wang, Sabrina Gaito, Haitao Zheng, and Ben Y Zhao. “Multi-scale dynamics in a massive online social network”. In: *Proceedings of the 2012 Internet Measurement Conference*. 2012, pp. 171–184.
- [11] Fariba Karimi and Petter Holme. “Threshold model of cascades in empirical temporal networks”. In: *Physica A: Statistical Mechanics and its Applications* 392.16 (2013), pp. 3476–3483. ISSN: 0378-4371. DOI: <https://doi.org/10.1016/j.physa.2013.03.050>. URL: <https://www.sciencedirect.com/science/article/pii/S0378437113002835>.
- [12] Naoki Masuda and Petter Holme. “Predicting and controlling infectious disease epidemics using temporal networks”. In: *F1000prime reports* 5 (2013).
- [13] Kosuke Yoshihara, Maria Shahmoradgoli, Emmanuel Martínez, Rahulsimham Vegesna, Hoon Kim, Wandaliz Torres-Garcia, Victor Treviño, Hui Shen, Peter W Laird, Douglas A Levine, et al. “Inferring tumour purity and stromal and immune cell admixture from expression data”. In: *Nature communications* 4.1 (2013), p. 2612.
- [14] Antoine Moinet, Michele Starnini, and Romualdo Pastor-Satorras. “Burstiness and aging in social temporal networks”. In: *Physical review letters* 114.10 (2015), p. 108701.

- [15] Avanti Athreya, Donniell E Fishkind, Minh Tang, Carey E Priebe, Youngser Park, Joshua T Vogelstein, Keith Levin, Vince Lyzinski, and Yichen Qin. “Statistical inference on random dot product graphs: a survey”. In: *The Journal of Machine Learning Research* 18.1 (2017), pp. 8393–8484.
- [16] Jose Caballero, Christian Ledig, Andrew Aitken, Alejandro Acosta, Johannes Totz, Zehan Wang, and Wenzhe Shi. “Real-time video super-resolution with spatio-temporal networks and motion compensation”. In: *Proceedings of the IEEE conference on computer vision and pattern recognition*. 2017, pp. 4778–4787.
- [17] Antoine Moinet, Romualdo Pastor-Satorras, and Alain Barrat. “Effect of risk perception on epidemic spreading in temporal networks”. In: *Phys. Rev. E* 97 (1 Jan. 2018), p. 012313. DOI: [10.1103/PhysRevE.97.012313](https://doi.org/10.1103/PhysRevE.97.012313). URL: <https://link.aps.org/doi/10.1103/PhysRevE.97.012313>.
- [18] Nikolaos Papachristou, Payam Barnaghi, Bruce Cooper, Kord M Kober, Roma Maguire, Steven M Paul, Marilyn Hammer, Fay Wright, Jo Armes, Eileen P Furlong, et al. “Network analysis of the multidimensional symptom experience of oncology”. In: *Scientific reports* 9.1 (2019), p. 2258.
- [19] Alba Contreras, Carmen Valiente, Alexandre Heeren, and Richard Bentall. “A temporal network approach to paranoia: A pilot study”. In: *Frontiers in psychology* 11 (2020), p. 544565.
- [20] Raj Dandekar and George Barbastathis. “Quantifying the effect of quarantine control in Covid-19 infectious spread using machine learning”. In: *medRxiv* (2020). DOI: [10.1101/2020.04.03.20052084](https://doi.org/10.1101/2020.04.03.20052084). eprint: <https://www.medrxiv.org/content/early/2020/04/06/2020.04.03.20052084.full.pdf>. URL: <https://www.medrxiv.org/content/early/2020/04/06/2020.04.03.20052084>.

- [21] D Gage Jordan, E Samuel Winer, and Taban Salem. “The current status of temporal network analysis for clinical science: Considerations as the paradigm shifts?” In: *Journal of clinical psychology* 76.9 (2020), pp. 1591–1612.
- [22] Christopher Rackauckas, Yingbo Ma, Julius Martensen, Collin Warner, Kirill Zubov, Rohit Supekar, Dominic Skinner, and Ali Jasim Ramadhan. “Universal Differential Equations for Scientific Machine Learning”. In: *CoRR* abs/2001.04385 (2020). arXiv: [2001.04385](https://arxiv.org/abs/2001.04385). URL: <https://arxiv.org/abs/2001.04385>.
- [23] Manuel A Roehrl, Thomas A Runkler, Veronika Brandtstetter, Michel Tokic, and Stefan Obermayer. “Modeling system dynamics with physics-informed neural networks based on lagrangian mechanics”. In: *IFAC-PapersOnLine* 53.2 (2020), pp. 9195–9200.
- [24] Junbo Zhang, Yu Zheng, Junkai Sun, and Dekang Qi. “Flow Prediction in Spatio-Temporal Networks Based on Multitask Deep Learning”. In: *IEEE Transactions on Knowledge and Data Engineering* 32.3 (2020), pp. 468–478. DOI: [10.1109/TKDE.2019.2891537](https://doi.org/10.1109/TKDE.2019.2891537).
- [25] Han Gao, Luning Sun, and Jian-Xun Wang. “PhyGeoNet: Physics-informed geometry-adaptive convolutional neural networks for solving parameterized steady-state PDEs on irregular domain”. In: *Journal of Computational Physics* 428 (2021), p. 110079. ISSN: 0021-9991. DOI: <https://doi.org/10.1016/j.jcp.2020.110079>. URL: <https://www.sciencedirect.com/science/article/pii/S0021999120308536>.
- [26] George Em Karniadakis, Ioannis G Kevrekidis, Lu Lu, Paris Perdikaris, Sifan Wang, and Liu Yang. “Physics-informed machine learning”. In: *Nature Reviews Physics* 3.6 (2021), pp. 422–440.

- [27] John A Keith, Valentin Vassilev-Galindo, Bingqing Cheng, Stefan Chmiela, Michael Gastegger, Klaus-Robert Müller, and Alexandre Tkatchenko. “Combining machine learning and computational chemistry for predictive insights into chemical systems”. In: *Chemical reviews* 121.16 (2021), pp. 9816–9872.
- [28] Aditi Krishnapriyan, Amir Gholami, Shandian Zhe, Robert Kirby, and Michael W Mahoney. “Characterizing possible failure modes in physics-informed neural networks”. In: *Advances in Neural Information Processing Systems* 34 (2021), pp. 26548–26560.
- [29] Maxime Lucas, Arthur Morris, Alex Townsend-Teague, Laurent Tichit, Bianca Habermann, and Alain Barrat. “Inferring cell cycle phases from a partially temporal network of protein interactions”. In: *Cell Reports Methods* (2021).
- [30] Mohammadamin Mahmoudabadbozchelou, Marco Caggioni, Setareh Shahsavari, William H Hartt, George Em Karniadakis, and Safa Jamali. “Data-driven physics-informed constitutive metamodeling of complex fluids: A multifidelity neural network (MFNN) framework”. In: *Journal of Rheology* 65.2 (2021), pp. 179–198.
- [31] Francesco Sanna Passino, Anna S Bertiger, Joshua C Neil, and Nicholas A Heard. “Link prediction in dynamic networks using random dot product graphs”. In: *Data Mining and Knowledge Discovery* 35.5 (2021), pp. 2168–2199.
- [32] Elaheh Kalantari, Samaneh Kouchaki, Christine Miaskowski, Kord Kober, and Payam Barnaghi. “Network analysis to identify symptoms clusters and temporal interconnections in oncology patients”. In: *Scientific Reports* 12.1 (2022), p. 17052.
- [33] Patrick Kidger. “On neural differential equations”. In: *arXiv preprint arXiv:2202.02435* (2022).

- [34] Thi Nguyen Khoa Nguyen, Thibault Dairay, Raphaël Meunier, and Mathilde Mougeot. “Physics-informed neural networks for non-Newtonian fluid thermo-mechanical problems: an application to rubber calendering process”. In: *arXiv preprint arXiv:2201.13389* (2022).