

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. QUESTIONS

Worth having definitions in appendices

3. LITERATURE REVIEW

3.1. **Temporal Networks.** Temporal networks are used to describe a series of interactions across time; an example of this can be found in XXX ref temp net example XXX. This can be represented in many ways, for example as a series of interactions or as series of adjacency matrices ordered by time. The series of interactions is generally represented as a list of tuples containing the node that interacted and the time that the interaction took place. In the representation of a series of adjacency matrices, the columns and rows of the matrix each refer to each of the nodes. In this

way, if an edge exists between nodes i and j , the entry (i, j) of the adjacency matrix will be 1; if an edge does not exist the entry will be 0.

Temporal networks are present in many areas of research interest, ranging from symptom interactions in mental health[19, 17], to epidemiology[12], protein interactions[26, 6], and social networks[14, 8].

Generally attempts to model temporal networks either model the change in node state, that is, some information about that node; for example the severity of a symptom at a given time, as in [17]. Or they model the changing structure of the network as in [28].

Contreras et al. [17] model the change in node state over time while the structure of the network remains constant. Contreras et al. look at the severity of symptoms over time, i.e. the state of the nodes, as a temporal network where the network structure remains constant. They use multilevel vector autoregression, a method that finds correlations between the current state of a variable and the states of variables at previous time steps [3], to model how the current symptoms of people with paranoia might predict their future symptoms. These models were then used to create three networks that linked symptoms. Of particular interest to this thesis is the temporal network created by having a fully connected network of all symptoms including self-loops, where a symptom connects to itself; this self loop comes about from a symptom being correlated with itself from one time step to the next. 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 be used to extrapolate and predict into the future.

In [11], the Karimi and Holme look to model the change in node state while the network structure changes. Karimi and Holme represent their temporal network as

a series of interactions (a list of tuples in the form (i, j, t) , where i, j are the nodes interacting and t is the time at which the interaction occurred). The authors use a model to predict when a node will change from state 0 to state 1, in practice this could represent whether a user of a social network believes a rumour, or an ecological patch being colonised. Notably, in their model the node never changes back; this is done to keep the model analytically tractable, but does limit its usefulness in some contexts. In this model an organism could not become locally extinct in an ecological patch for example. To model the change in a nodes' state, the model uses a sliding time window and if the fraction of interactions with nodes of state 1 within that time window exceeds a given threshold, the node switches to state 1 as well.

This paper does not aim to present a model for predicting the state of a node in the future. Instead, it aims to present a tool for understanding the spread of ideas, viruses, etc. within a network.

Pasino et al.[28] model the change in the structure of a temporal network. In this case the Sanna Passino et al. 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 using a spectral embedding (SVD). A spectral embedding uses the eigenvectors of a network's adjacency matrix to find vectors that represent the nodes of the network. The spectral embedding transforms their temporal sequence into a sequence of observations of a latent space. A latent space is a continuous space in which similar nodes (nodes with many of the same neighbours) are positioned near each other. In this space nodes are represented as points in a continuous, low dimensional space. Another feature of this latent space, is that the dot product of any two points in the space gives us the probability that the nodes related to those points are connected. With this series of embeddings, the authors employ a variety of time series techniques to predict the network structure at future time steps.

3.2. UDEs and NNDEs. In some temporal 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 neighbours[10, 5]. We also see that the state of symptoms (occurrence, severity, and distress), can influence the states of other symptoms in patients undergoing chemotherapy [16, 29], and for mental health disorders [17]. As well as in ecological networks where the interactions of species are often represented as a network, and the populations of one species may influence the population growth of another [2, 1]; that is, the state (in this case the population) of a species (node), may influence the movement of another species.

With the observation that nodes in real world temporal networks may influence others, differential equations or difference equations seem appropriate for modelling the evolution of nodes as there is a rich body of literature exploring these methods use in modelling interacting variables. 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. We note however, that the progression of the network is generally continuous even if the events happen as discrete jumps. Because of this, we may wish to preserve the continuous temporal progression in our model.

Looking at the work of Passino et al.[28], the authors use random dot produce 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. That is, we will be able to model the changes in the probabilities of edges continuous and from this, generate a network at any time step.

Using the method in Passino et al.[28], we can treat the sequence of embeddings as a dynamical system and then use differential equation modelling in a continuous space. That is, once we have found a latent space for each network. We know that each point in this space refers to a specific node in the network, and so, as the network changes, so too does the position of each of the points in the latent space. We then look to model this movement using differential equations. Our aim for this framework to be as general as possible and note that there may not yet be enough domain knowledge to perform traditional differential equation modelling to the desired accuracy. For example in network ecology models can be useful, but not perfect XXX cite XXX. Hence, we might wish to build on these models rather than trying to create something entirely new. This can be achieved by using this partial domain knowledge using it to create a universal differential equation (UDE) [20].

3.2.1. *Theory.* Developed by Rackauckas et al.[20], Universal differential equation (UDEs) are a novel neural network architecture that aims to take the best of both traditional neural differential equation 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. This hybrid machine learning model is then trained on the observed data. Combining the two allows for most of the movement in the data to be captured by the differential equation, which contains the domain knowledge. In this way the neural network will only need to learn a theoretically simpler equation, given that a part of the system has already been captured. This then allows the neural network to be smaller, require fewer 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 entirely understood[30]. N

There is an ecosystem of packages specifically designed for performing machine learning with UDEs. The ecosystem is called SciML or scientific machine learning[20]. 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.

3.2.2. Applications. 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) [23, 22, 25, 21], and on improving modelling of fluids [27, 31]. Alongside 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 [18]. 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 networks. This seems to be a natural fit for UDEs and neural ODEs, given we may know relatively little about the processes that govern the evolution of temporal networks, and any information we do have can be incorporated into the model.

3.3. Symbolic Regression. Symbolic regression is a type of regression analysis that searches a space of equations to find an expression that matches data. This can be achieved in many ways and remains active area of research. One of the most popular algorithms is genetic programming[7]. Genetic programming has been the subject of extensive research, and various improvements to speed and efficiency have been made XXX cite XXX. Genetic programming involves constructing a search tree using a genetic algorithm to yield a potential equation, which may match the data as well as numerically calculated partial derivatives of each of the input variables. Comparing with these derivatives is useful as it helps to ensure that the produced

function equations have some grounding in the real world process that generated the data. The search tree is then pruned by comparing the partial derivatives of the candidate equation, with partial derivatives calculated from data. This process is then iterated many times to generate a suitable equation.

The problem that symbolic regression aims to solve is to generate interpretable and “meaningful” equations from observed data. For example, both [7, 4] test their methods’ capacity for obtaining a symbolic equation which physical simulation data, various pendula specifically. A “meaningful” equation in this case would be one that adheres to the laws of physics.

Examples of methods include discovering equations governing the movement of simple harmonic and chaotic double pendula from data[7]. 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. In [Cite rackaucas presentation], [i- authors] demonstrate that using symbolic regression, as opposed to just an NNDE, can improve the accuracy of extrapolating. One aspect of this improvement is having accurate numerical derivative to compare to. To ensure this, it is important that the data is sampled very frequently. This seems integral to the accuracy of the found equation.

Another approach to symbolic regression can be found in [4]. 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 predictions are maximised. The algorithm will generate new initial conditions to find behaviour exhibited by the numerical equation, for example a fixed point, that is not exhibited by the functional equation. When a difference is found, the functional equation is adjusted to also exhibit the behaviour of the numerical equation. 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 the scope of the training data.

Having interpretable models that capture underlying relations is relevant to all fields of study, and so the applications of symbolic regression are incredibly varied. Various methods of symbolic regression have been use in manufacturing systems, chemical systems, and tumor research [9, 24, 13]. As well research into improvements is still on going.

In very recent work, Kidger[30], demonstrates that, with NNDEs, two of the assumptions necessary for symbolic regression can be overcome. Those are the necessity for paired observations and derivatives, as well as the assumption that the function can be expressed as a shallow tree of symbolic operations. We are modelling a differential equation, and so will have access to the derivatives at any point.

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] Judith D Singer, John B Willett, John B Willett, et al. *Applied longitudinal data analysis: Modeling change and event occurrence*. Oxford university press, 2003.
- [4] 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.
- [5] 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.

- [6] 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.
- [7] Michael Schmidt and Hod Lipson. “Distilling free-form natural laws from experimental data”. In: *science* 324.5923 (2009), pp. 81–85.
- [8] Steve Hanneke, Wenjie Fu, and Eric P Xing. “Discrete temporal models of social networks”. In: (2010).
- [9] 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.
- [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] 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.
- [17] 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.
- [18] 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>.
- [19] 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.
- [20] 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>.

- [21] 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.
- [22] 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>.
- [23] 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.
- [24] 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.
- [25] 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.
- [26] 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).
- [27] 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.
- [28] 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.
- [29] 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.
- [30] Patrick Kidger. “On neural differential equations”. In: *arXiv preprint arXiv:2202.02435* (2022).
- [31] 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).