Deep Learning Based Time Evolution

Geoffrey Fox, Indiana University

Abstract
We show that one can study several time-series in terms of an underlying time evolution operator which can be learned with a recurrent deep learning network. This has been shown for Newton's laws for particles and Covid case and death data from observation and models while other work has studied this successfully in transportation systems. We propose to extend this research to full epidemiological simulations, earthquake forecasting (in progress) and networking and compare the successful deep learning architectures in each case to understand how application characteristics map into the most successful deep learning structures considering recurrent, convolutional, graph and fully connected linkages as well as sequence to sequence mapping approaches such as the transformer network. The role of spatial structure and multiple time scales and hierarchical deep networks will be considered.

Introduction
There is increasing recognition of the importance of deep learning in data-driven discovery across a broad range of applications.Here we study time series where the MLPerf [1], [2] time series working group has recently highlighted many areas and available datasets [3]. Logistics, network intelligence, manufacturing, smart city, and ride-hailing [4] (transportation) are major Industry areas having important time series while medical data is often of this form. We note that similar technical approaches (recurrent neural nets) are often used for both time series and "sequence to sequence mapping" as seen in the major voice and translation areas separately studied at MLPerf. We focus here on the analysis of time-dependent data where our approach can be illustrated by the three examples below

Deep Learning as a Particle Dynamics Integrator

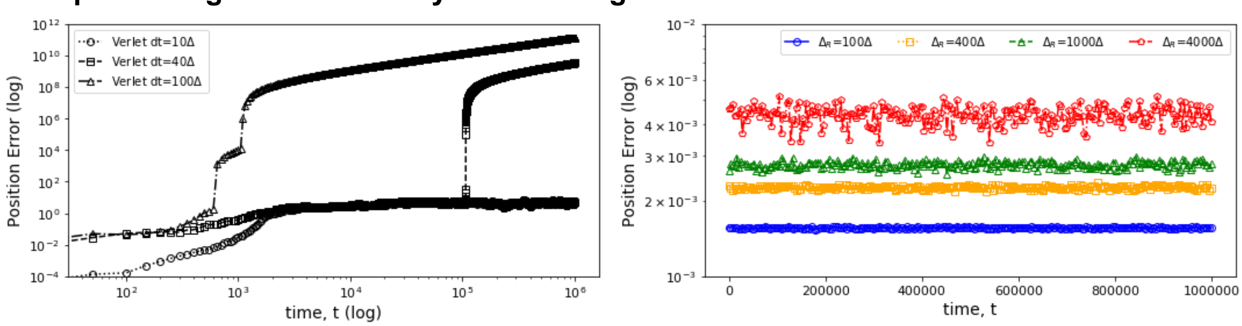

Fig. 1. The average error in position updates for 16 particles interacting with an LJ potential, The left figure is standard MD with error increasing for ∆t as 10, 40, or 100 times robust choice (0.001). On the right is the LSTM network with modest error up to t = 10^6 even for ∆t = 4000 times the robust MD choice.

Molecular dynamics simulations rely on numerical integrators to solve Newton's equations of motion. Using a sufficiently small time step to avoid discretization errors, these integrators generate a trajectory of particle positions as solutions to the equations of motions. In [5]–[7], the IU team introduces an integrator based on recurrent neural networks that is trained on trajectories generated using the Verlet integrator and learns to propagate the dynamics of particles with timestep up to 4000 times larger compared to the Verlet timestep. As shown in fig.

1 (right) the error does not increase as one evolves the system for the surrogate while the standard integration in fig. 1 (left) has unacceptable errors even for time steps of just 10 times that used in an accurate simulation. The surrogate demonstrates a significant net speedup over Verlet of up to 32000 for few-particle (1 - 16) 3D systems and over a variety of force fields including the Lennard-Jones (LJ) potential.

We often think of the laws of physics described by operators that evolve the system given sufficient initial conditions and in this language, we have shown how to represent Newton's law operator by a recurrent network. We expect that the time dependence of many complex systems: Covid pandemics, Southern California earthquakes, traffic flow, security events can be described by deep learning operators that both capture the dynamics and allow predictions. In the covid example below for example one can learn an operator that depends on the demographics and social distancing approach for a given region.

Deep Learning to describe Covid Daily Data

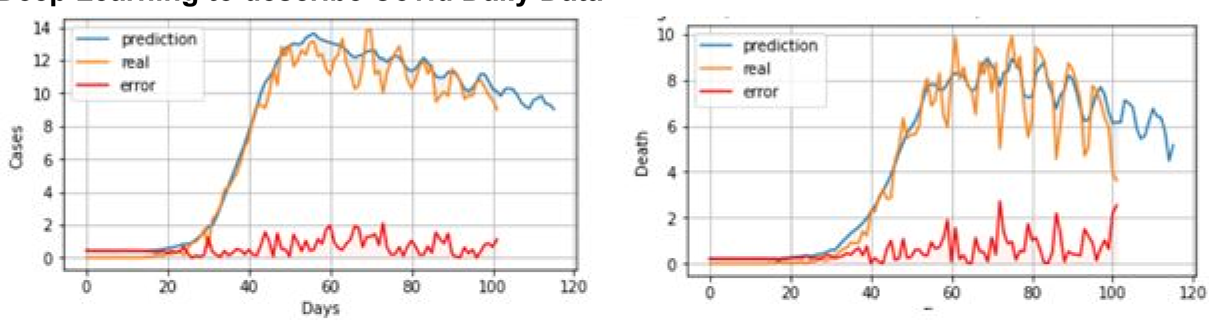

Fig 2: Deep Learning fits to Covid case and death data from Feb. 1 to May 25, 2020, with predictions 2 weeks out and showing a weekly structure

There are extensive collections of daily data for the number of Covid reported cases and deaths. These can be described by epidemiological models plus empirical fits [8] but as proposed above and illustrated in fig. 2, we developed a deep learning model [9] that learned a Covid daily evolution operator from 110 separate time series of curated (by the University of Pittsburgh) data for different US cities. The time series were 100 days long and the model was a 2 layer LSTM recurrent network similar to that used to describe the evolution of molecular dynamics above. It differed by learning from the demographics (fixed data for each city) as well as time-dependent data and by predicting ahead for two weeks with each series as shown in the figure. This capability is important in any application with multiple time scales. For example, in earthquake forecasting multiscale in time effects are critical and one might want to combine a general forecast for the next time step (days to months) with the probability of the big one happening in the next 10 years. For 37 of the 110 cities reliable empirical (not deep learning) fits are available to the case and death data up to April 15, 2020 [8]. A single deep learning time evolution operator can describe these 37 separate datasets and smooth fitted data leads to very accurate deep learning descriptions shown in fig. 3. For both figs. 2 and 3, the data is divided into windows of size 5, 9, or 13, and cases and deaths were simultaneously trained together with demographic data. This surrogate for an empirical fit will be generalized to a surrogate for a sophisticated epidemiological simulation. We will also need to link with time-dependent mobility and social distancing data[10].

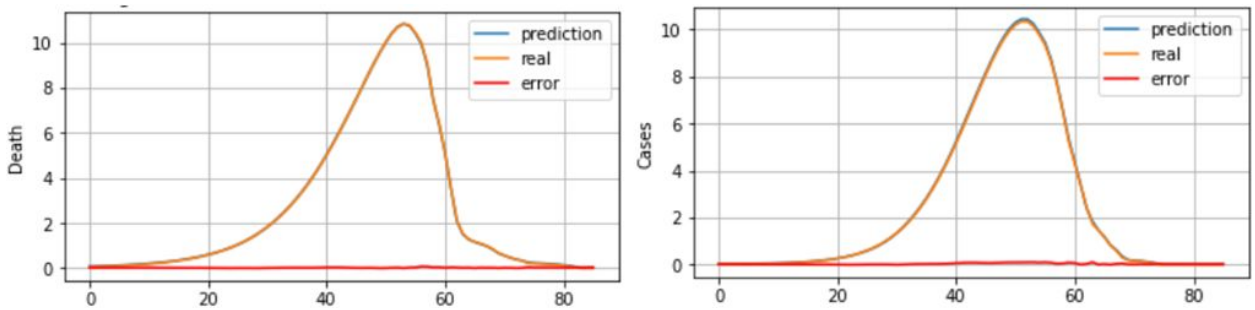

Fig 3: Deep Learning Fits empirical Covid data descriptions with 37 separate results shown as summed over cities. The cases and death were learned together in time series for different locations

Above we have given 3 examples of recurrent networks of the time evolution operator for complex systems and we are extending this to other areas. We see the mix of dense and recurrent networks used above as a base approach applicable to many problems. Some examples need additional features: earthquakes (with fault lines) and transportation (road systems) need graph networks while mixtures of convolutional and recurrent networks (such as convLSTM) are used in weather and again earthquakes where the time series features can consist of images. We intend to study deep learning based time evolution operators for different complex systems and identify patterns as to which type of network describes which problem classes and the amount of data needed to get good results. Hopefully we will also make research advances in the best networks to use; this is already seen in the move from recurrent networks to transformer and reformer architectures but this was largely motivated by sequence to sequence mapping and not by time series. We suggest more research in multiple or hierarchical time scales as this is needed in many applications.

We see this collection of time series datasets and reference implementations as playing the same role for time series that ImageNet ILSRVC and AlexNet played for images. The different implementations establish best practice, get chosen for different application areas to either suggest an architecture or an initial network by transfer learning. Interesting complex systems that we can quickly look at include virtual tissues [11], [12] and epidemiology[13] for Covid related applications. Such evolution operators are also seen[3] in finance, networking, security, monitoring of complex systems from Tokomaks [14] to operating systems, and environmental science.

Acknowledgements
This work is partially supported by the National Science Foundation (NSF) through awards CIF21 DIBBS 1443054, nanoBIO 1720625, CINES 1835598 and Global Pervasive Computational Epidemiology 1918626. I thank Gregor von Laszewski, Saumyadipta Pyne, JCS Kadupitiya, and Vikram Jadhao for great discussions.

References

[1] P. Mattson, C. Cheng, G. Diamos, C. Coleman, P. Micikevicius, D. Patterson, H. Tang, G.-Y. Wei, P. Bailis, V. Bittorf, D. Brooks, D. Chen, D. Dutta, U. Gupta, K. Hazelwood, *et al.*, “MLPerf Training Benchmark,” in *Proceedings of Machine Learning and Systems 2020*, 2020, pp. 336–349.

[2] "MLPERF benchmark suite for measuring performance of ML software frameworks, ML hardware accelerators, and ML cloud platforms." [Online]. Available: https://mlperf.org/. [Accessed: 08-Feb-2019]
[3] X. Huang, G. C. Fox, S. Serebryakov, A. Mohan, P. Morkisz, and D. Dutta, "Benchmarking Deep Learning for Time Series: Challenges and Directions," in *2019 IEEE International Conference on Big Data (Big Data)*, 2019, pp. 5679–5682 [Online]. Available: http://dx.doi.org/10.1109/BigData47090.2019.9005496
[4] Yan Liu, "Artificial Intelligence for Smart Transportation Video." [Online]. Available: https://slideslive.com/38917699/artificial-intelligence-for-smart-transportation. [Accessed: 08-Aug-2019]
[5] JCS Kadupitiya, Geoffrey C. Fox, Vikram Jadhao, "GitHub repository for Simulating Molecular Dynamics with Large Timesteps using Recurrent Neural Networks." [Online]. Available: https://github.com/softmaterialslab/RNN-MD. [Accessed: 01-May-2020]
[6] J. C. S. Kadupitiya, G. C. Fox, and V. Jadhao, "Simulating Molecular Dynamics with Large Timesteps using Recurrent Neural Networks," *arXiv [physics.comp-ph]*, 12-Apr-2020 [Online]. Available: http://arxiv.org/abs/2004.06493
[7] J. C. S. Kadupitiya, G. Fox, and V. Jadhao, "Recurrent Neural Networks Based Integrators for Molecular Dynamics Simulations," in *APS March Meeting 2020*, 2020 [Online]. Available: http://meetings.aps.org/Meeting/MAR20/Session/L45.2. [Accessed: 23-Feb-2020]
[8] Robert Marsland and Pankaj Mehta, "Data-driven modeling reveals a universal dynamic underlying the COVID-19 pandemic under social distancing," *arXiv [q-bio.PE]*, 21-Apr-2020 [Online]. Available: http://arxiv.org/abs/2004.10666
[9] Luca Magri and Nguyen Anh Khoa Doan, "First-principles Machine Learning for COVID-19 Modeling," *Siam News*, vol. 53, no. 5, Jun. 2020 [Online]. Available: https://sinews.siam.org/Details-Page/first-principles-machine-learning-for-covid-19-modelin g
[10] A. Adiga, L. Wang, A. Sadilek, A. Tendulkar, S. Venkatramanan, A. Vullikanti, G. Aggarwal, A. Talekar, X. Ben, J. Chen, B. Lewis, S. Swarup, M. Tambe, and M. Marathe, "Interplay of global multi-scale human mobility, social distancing, government interventions, and COVID-19 dynamics," *medRxiv - Public and Global Health*, 07-Jun-2020 [Online]. Available: http://dx.doi.org/10.1101/2020.06.05.20123760
[11] T. J. Sego, J. O. Aponte-Serrano, J. F. Gianlupi, S. Heaps, K. Breithaupt, L. Brusch, J. M. Osborne, E. M. Quardokus, and J. A. Glazier, "A Modular Framework for Multiscale Spatial Modeling of Viral Infection and Immune Response in Epithelial Tissue," *BioRxiv*, 2020 [Online]. Available: https://www.biorxiv.org/content/10.1101/2020.04.27.064139v2.abstract
[12] Y. Wang, G. An, A. Becker, C. Cockrell, N. Collier, and M. Craig, "Rapid community-driven development of a SARS-CoV-2 tissue simulator," *BioRxiv*, 2020 [Online]. Available: https://www.biorxiv.org/content/10.1101/2020.04.02.019075v2.abstract
[13] D. Machi, P. Bhattacharya, S. Hoops, J. Chen, H. Mortveit, S. Venkatramanan, B. Lewis, M. Wilson, A. Fadikar, T. Maiden, C. L. Barrett, and M. V. Marathe, "Scalable Epidemiological Workflows to Support COVID-19 Planning and Response," May 2020.
[14] J. Kates-Harbeck, A. Svyatkovskiy, and W. Tang, "Predicting disruptive instabilities in controlled fusion plasmas through deep learning," *Nature*, vol. 568, no. 7753, pp. 526–531, Apr. 2019 [Online]. Available: https://doi.org/10.1038/s41586-019-1116-4