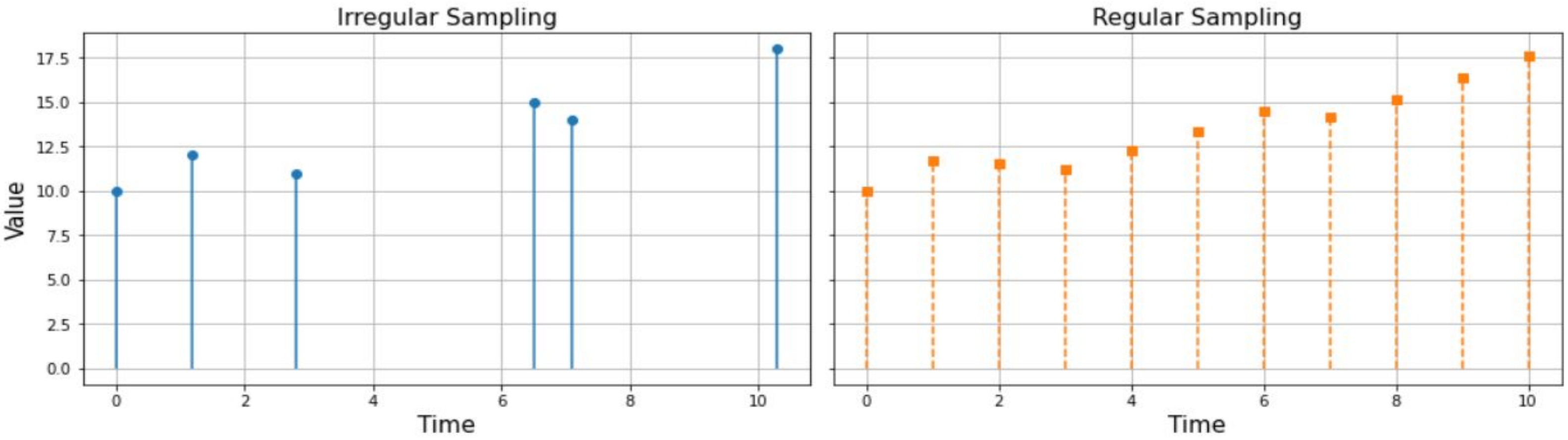
Neural Wave Equation for Irregularly Sampled Sequence Data

Arkaprava Majumdar,M Anand Krishna,Srijith PK

భారతీయ సాంకేతిక విజ్ఞాన సంస్థ హైదరాబాద్
भारतीय प्रौद्योगिकी संस्थान हैदराबाद
Indian Institute of Technology Hyderabad

BRAIN

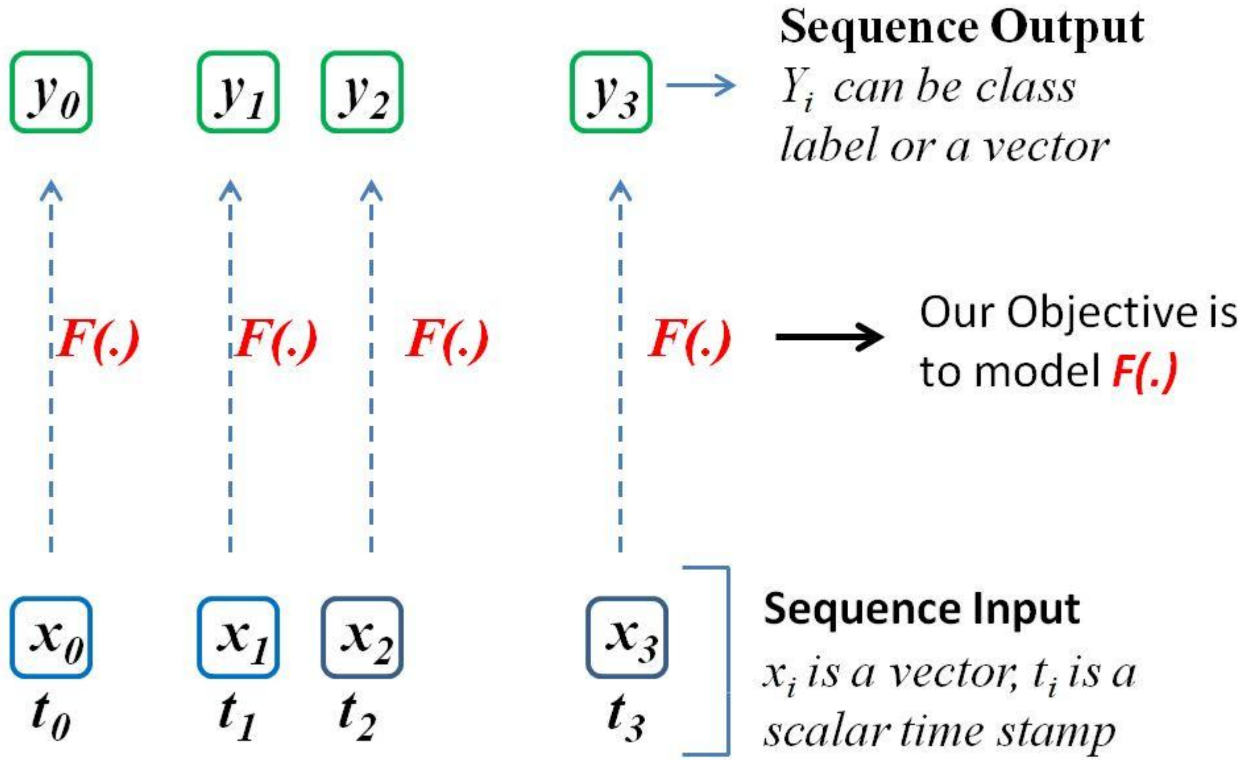
Real life examples of Irregularly sampled data



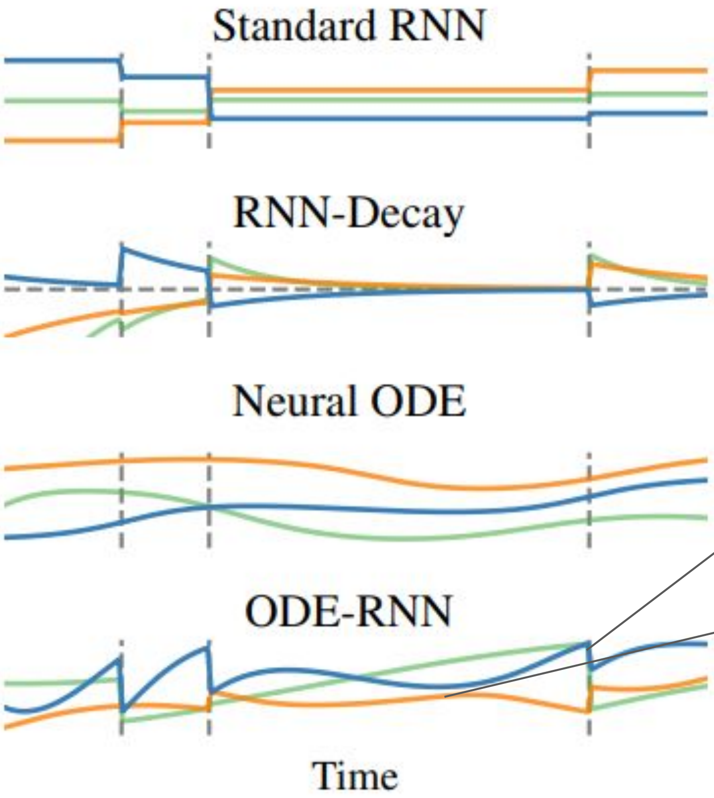
- Electronic Health Records
- Social Media Posts
- ClickStream

- Sensor Data
- Financial Data
- Medical Monitoring
- Weather models

Problem Formulation



ODE - RNN - the current standard for modelling irregular sequence data



$$h_k^d = \sigma(W_{\bar{h}} h_k^{d-1} + W_h \hat{h}_k^d + b_h)$$

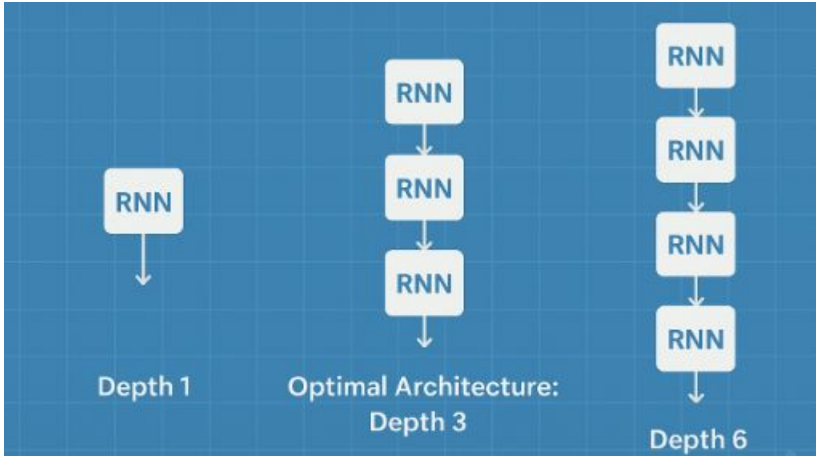
$$\hat{h}_k^d = ODESolve(h_{k-1}^d, (t_{k-1}, t_k))$$

Time points when a new input is received. We model this jump by a RNNCell

Continuous trajectory in between two inputs is modelled by an ODE

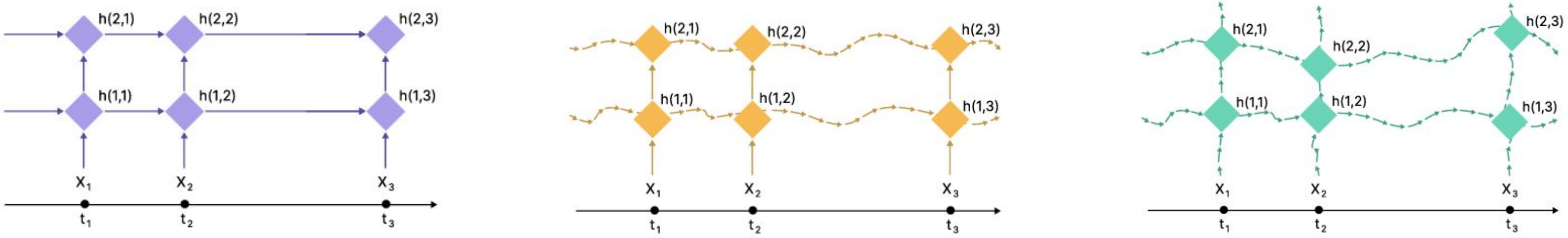
Issues with ODE-RNN - The depth selection

- ODE-RNN aims to model the hidden state in between two inputs using a neural ODE.
- However, it is still discrete in the depth direction.
- From a preliminary experiment on electrical transformers data, we notice that the accuracy of a model depends heavily on the depth of the model.



RNN and ODE-RNN models require careful depth selection, making the model selection phase both important and resource-intensive.

Modelling both time and depth continuously



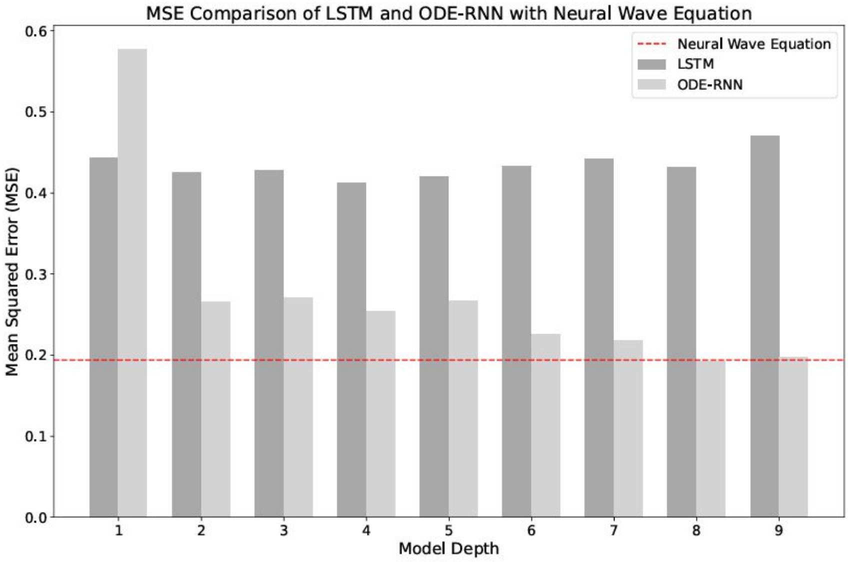
A - RNN Variant B - ODE-RNN Variant C - Neural Wave Equation

- Our method alleviates the need for depth selection by modelling both depth and time continuously using a non-homogeneous wave equation.
- In a way, we extend the original neural ODE's philosophy of automatic depth selection via an adaptive step size solver in sequence to sequence problems.

Empirical Results on why depth selection is crucial

- Our results demonstrate that ODE-RNN architectures exhibit significant variance in performance across different depths. This highlights the importance of depth tuning and suggests that integrating automatic depth selection mechanisms could make ODE-RNNs more robust in practice.

- On the other hand, Neural Wave Equation, does not suffer from the model selection problem and outperforms the ODE-RNN models.



Neural Wave Equations

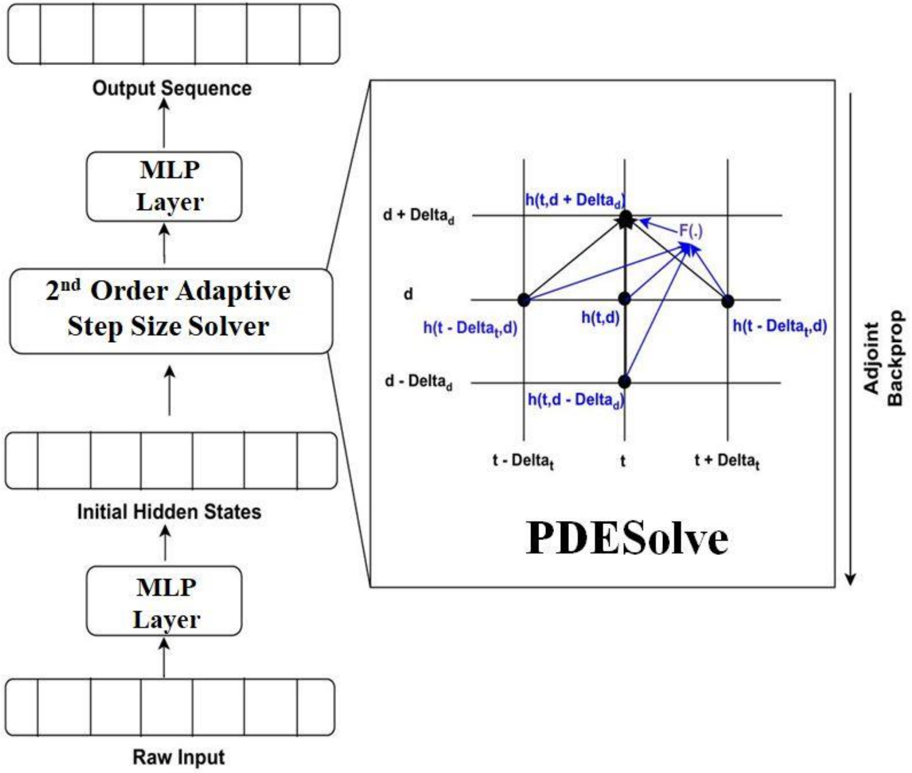
- We model both depth and time continuously using a partial differential equation.
- We majorly focus on using the wave equations in our work.

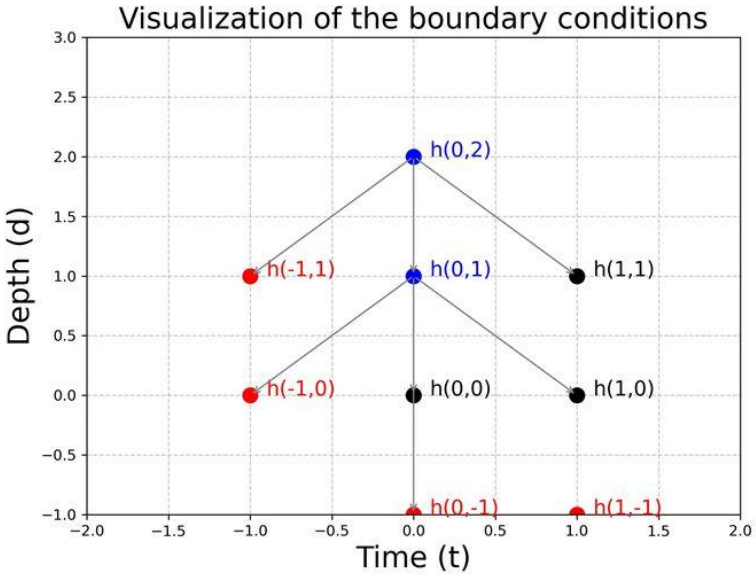
$$\frac{\partial^2 h_{t,d}}{\partial d^2} - c^2 \frac{\partial^2 h_{t,d}}{\partial t^2} = F_{\theta_s}(h_{t,d}, h_{t-\Delta_t,d}, h_{t+\Delta_t,d}, h_{t,d-\Delta d})$$

$$h_{t,d+\Delta_d} = 2h_{t,d} - h_{t,d-\Delta_d} + \frac{\Delta_d^2}{\Delta_t^2} c^2 [h_{t+\Delta t,d} - 2h_{t,d} + h_{t-\Delta_t,d}] + F_{\theta_s}(h_{t-\Delta_t,d}, h_{t,d}, h_{t+\Delta_t,d}, h_{t,d-\Delta_d})$$

The F(.) term models non-linear interaction between hidden states both across time and depth.

Architectural Details and Boundary Condition





Why PDEs work for sequence labelling

Aggregates information through depth

Wave Equation

Aggregates information in temporal direction

$$h(t,d) = \frac{f(t+cd)+f(t-cd)}{2} + \frac{1}{2c}\int_0^d \int_{t-c(d-\tau)}^{t+c(d-\tau)} F_{\theta_s}(h_{s,\tau}, h_{s-\Delta_s,\tau}, h_{s+\Delta_s,\tau}, h_{s,\tau-\Delta_\tau})\, ds\, d\tau$$

Heat Equation

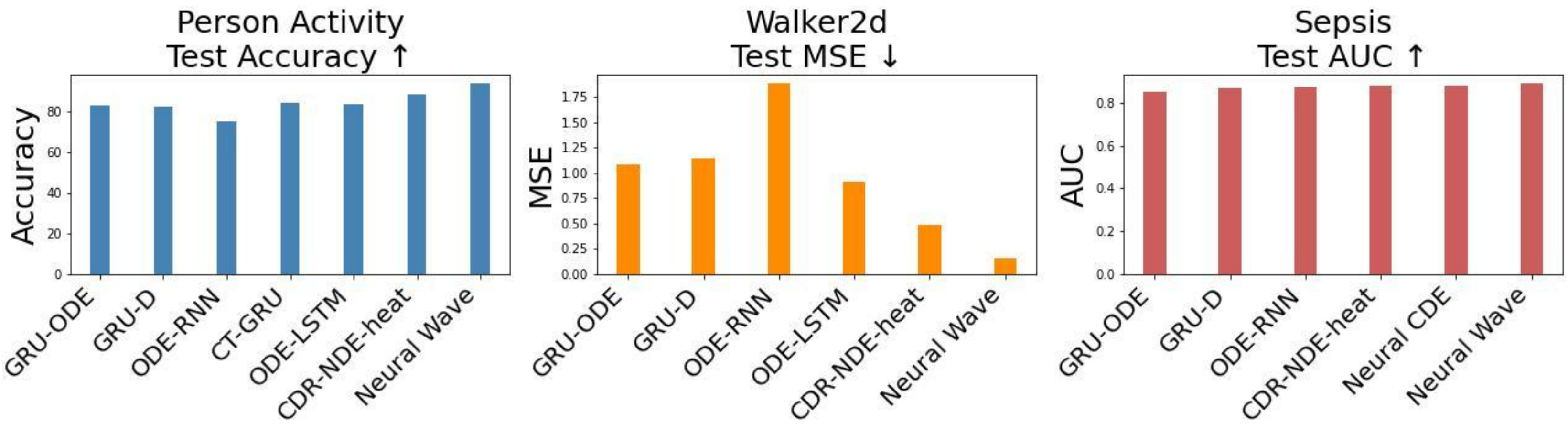
$$h(t,d) = \sum_{n=1}^{\infty} (a_n(0)\exp^{(-k\lambda_n d)} + \int_0^d q_n(\tau)\underline{\exp^{(-k\lambda_n(t-\tau)}\, d\tau})) \phi_n(t)$$

Even though, similar to wave equation, the information is exponentially decayed due to the underlined term

Normal RNN

$$h_{t,d} = F(W_t h_{t-1,d} + W_{d-1} h_{t,d-1})$$

Evaluation on standard benchmarks



Our results demonstrate that the Neural Wave Equation achieves superior performance compared to continuous-time RNN baselines on two key irregular sequence modeling benchmarks: the Person Activity and Sepsis Prediction datasets.

Conclusion and future remarks

- Partial Differential Equations can model time series continuously across depth and time.

- It alleviates the problem of depth selection in time series and outperforms other continuous time series models.

- In our setup, we experiment with heat and wave equations mainly because of their analytical properties. In future, experimenting with different PDEs may yield better results for certain types of time series data.