Constrained Dynamical Neural ODE for Time Series Modelling: A Case Study on Continuous Emotion Prediction

Ting Dang, Antoni Dimitriadis, Jingyao Wu, Vidhyasaharan Sethu, Eliathamby Ambikairajah
Contact: ting.dang@unimelb.edu.au

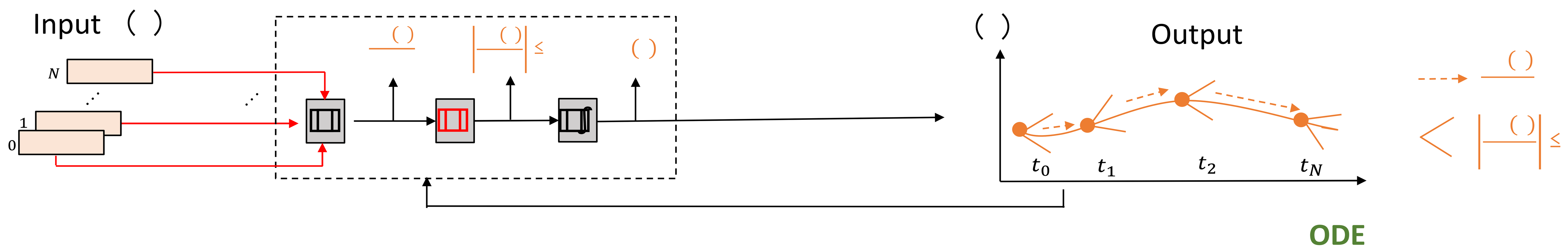
Highlights

*Leveraging known **constraints in time series modelling** via constrained dynamical neural ordinary differential equation (CD-NODE) can improve model accuracy with **fewer parameters** while achieving **faster convergence**.*

1 Motivations and Challenges

- Time series prediction is of great interest in various applications.
- The nature and characteristics of time series can vary significantly.
- For example, changes in emotion may be correlated with specific behavioural events, e.g., laughter; and heart rate changes relate to activities
- **Incorporating known constraints** on the changes for modelling may be beneficial.
- Existing models do not **expose the dynamics of the predicted quantity**, making it difficult to incorporate any prior knowledge

2 CD-NODEs

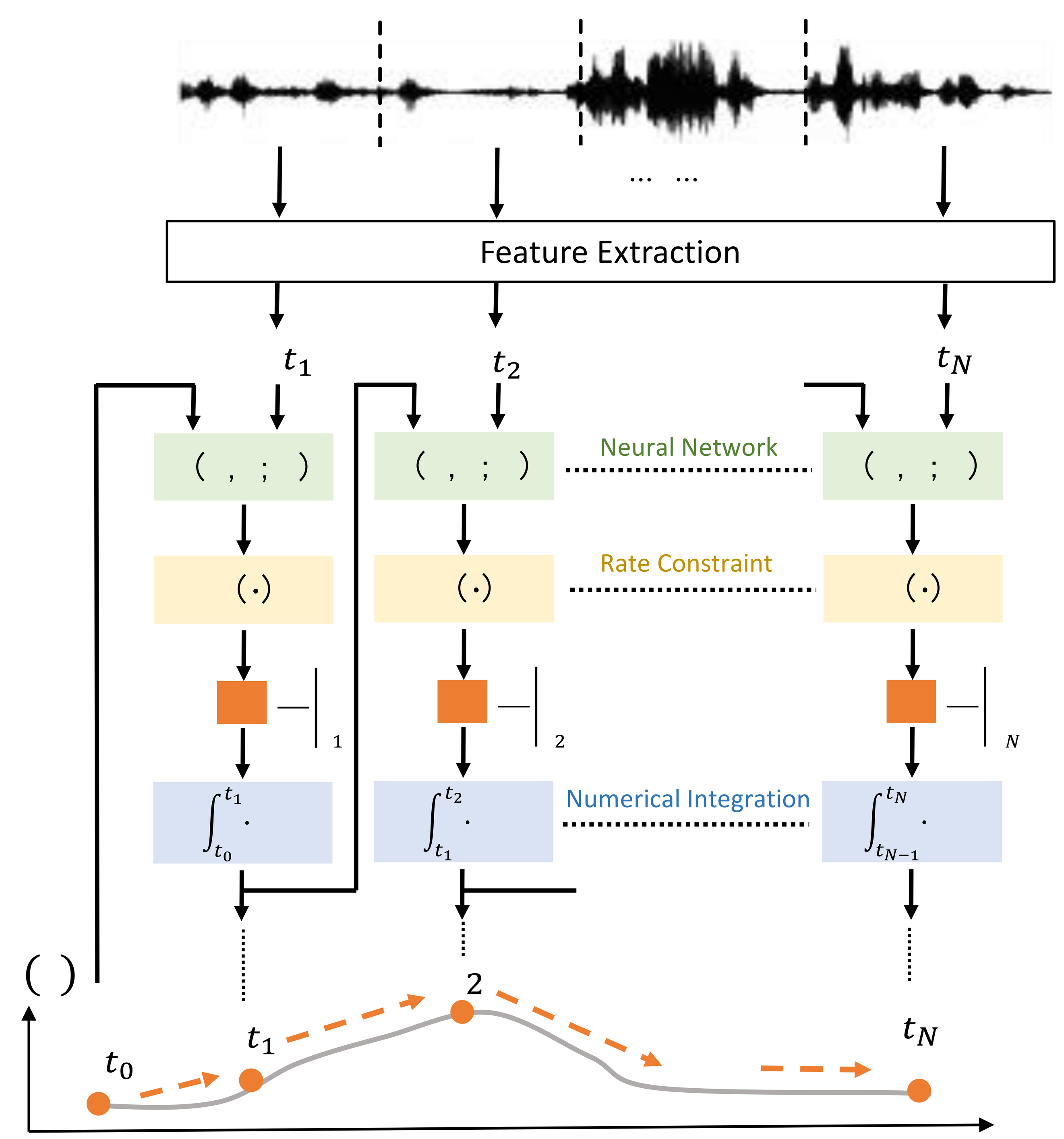


$$\frac{dy(t)}{dt} = \alpha * tanh\left(\frac{1}{\alpha} f(\boldsymbol{x}(t), y(t); \boldsymbol{\theta})\right)$$

- Input dependent dynamics: $\frac{dy(t)}{dt}$ is modelled as the input-driven governing function, varying with the input $x(t)$ and past $y(t)$.
- Rate constraint α: A large α would be akin to an unconstrained CD-NODE, while a small α limits the value of the derivatives

3 Models

Proposed CD-NODEs for Emotion



$$y(t) = ODESolve(f, \boldsymbol{\theta}, y(t_0), \boldsymbol{x}(t), t_n), \quad t_n \in [t_0, t_N]$$

- **Emotion changes are better perceived** when compared with numerical ratings of the emotional state.
- Bag-of-Audio-Word (**BoAW**) features $\boldsymbol{x}(\boldsymbol{t})$
- **Three fully connected layers** to approximate function f
- α values are selected within { 0.1, 0.25, 0.5, 1, 2, 4, 6, 8 } given **maximum rate of change as 6.25 and 3.88** for arousal and valence respectively in the training dataset.
- **Adjoint sensitivity method** used for backpropogation with loss:

$$L(T, \boldsymbol{\theta}) = 1 - CCC(y(t), \hat{y}(t))$$

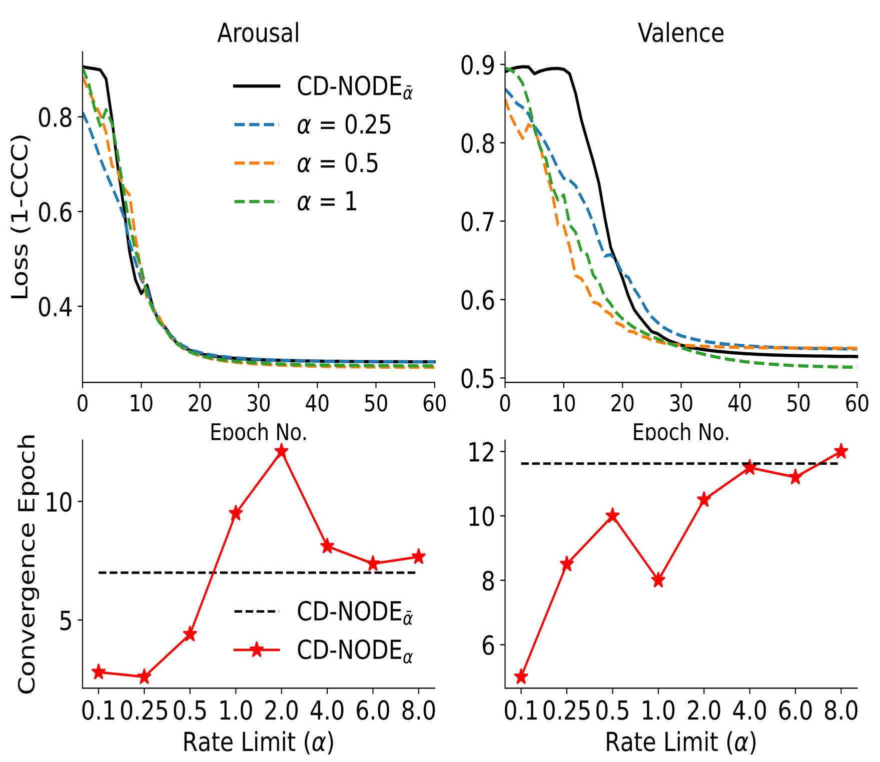
4 Results

Performance Comparison

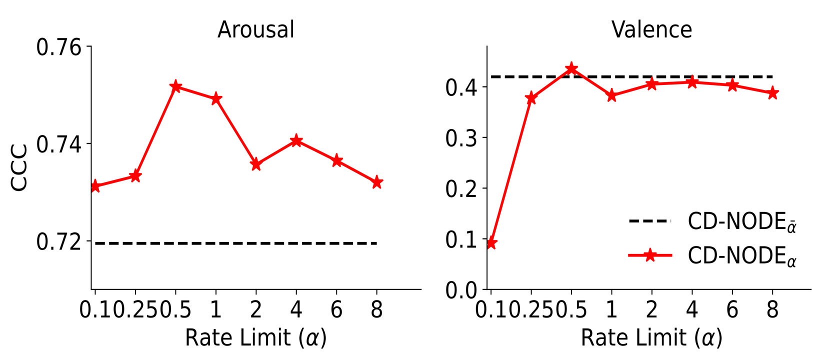
Systems	Features	Arousal	Valence
End-to-end	Raw signals	0.741	0.325
Adversarial	Functionals	**0.797**	0.474
Adversarial^{wd}	Functionals	0.780	**0.501**
Reconstruction	Functionals	0.754	0.378
LSTM	BoAW	0.728(0.098)	0.396(0.145)
CD-NODE$_{\bar{\alpha}}$	BoAW	**0.782**(0.052)	**0.506**(0.119)
CD-NODE$_{\alpha}$	BoAW	0.778(0.072)	0.491(0.115)

*wd: Wasserstein Distance used in adversarial training.

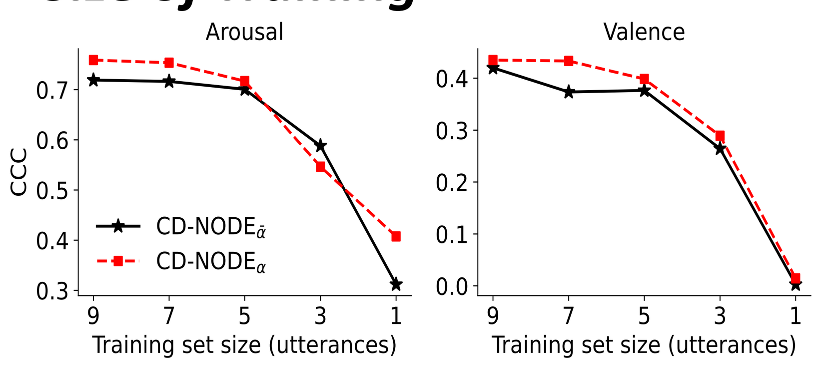
Convergence



Rate



Size of Training



5 Conclusion

- CD-NODEs allows for explicit constraints on dynamics of desired time series, e.g., input-driven nature and maximum rate of change.
- Modelling emotion dynamics with known constraints is more advantageous than directly modelling the numerical attributes.
- Rate constraint enables a faster convergence with fewer model parameters.

Code: https://github.com/TingDang90/CD-NODEs