Machine Learning and Data Intensive Computing (Mining) LAB

ICML
International Conference On Machine Learning

Deep Temporal Sets with Evidential Reinforced Attentions for Unique Behavioral Pattern Discovery

Dingrong Wang, Deep Shankar Pandey, Krishna Prasad Neupane, Zhiwei Yu, Ervine Zheng, Zhi Zheng, Qi Yu

Golisano College of Computing and Information Sciences
Rochester Institute of Technology (RIT)

Background and Challenges

We aim to perform **complex human behavioral analysis** and propose to model it as a temporal sequence prediction problem.

Key challenges involve:

- **Multimodal Pattern Extraction**: Human behavior data is inherently multimodal, where complex spatiotemporal patterns are hidden and hard to identify.
- **Robustness**: Temporal data may be incomplete and noisy, requiring robustness in partial prediction and temporal permutation invariance.
- **Interpretability**: Results of analysis may be used in mental health research and applications, which require a higher level of interpretability, making many black box models (e.g., deep neural networks) unsuitable.

We conduct comprehensive analysis on behaviors from children with and without Autism Spectrum Disorder (ASD) using the self-collected game (Maze Painting, Coloring, Word Scanning) datasets.

Summary of Contribution

- **Our model: Deep Temporal Set with Evidential Reinforced Attentions (DTS-ERA) aims to:**
  - **Extract complex spatiotemporal patterns from temporal data**
  - **Identify signature behavioral patterns (SBPs) from multimodal sequential data**
  - **Detect children with special behaviors (e.g., ASD) by leveraging discovered SBPs as key features.**

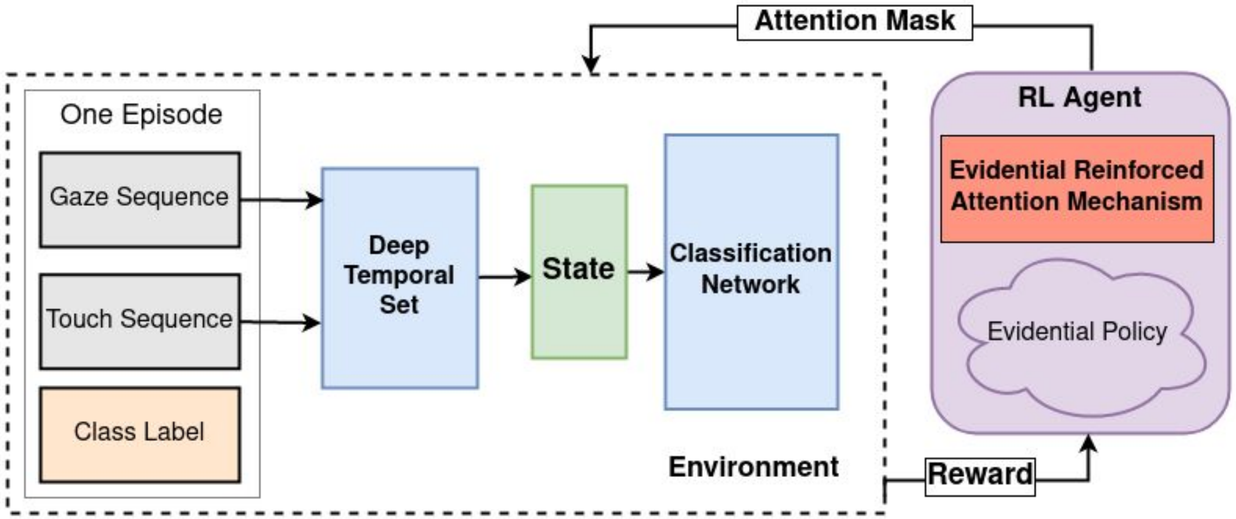
- **Develop a model $\mathcal{F}$ which satisfies:**

$$\mathcal{F}: \{\mathbf{g}_n, \mathbf{t}_n\}_{n=1}^{N_e} \rightarrow y;\ \mathbf{g}_n \in \mathbb{R}^{M_g}, \mathbf{t}_n \in \mathbb{R}^{M_t} \quad y \in [0,1]$$

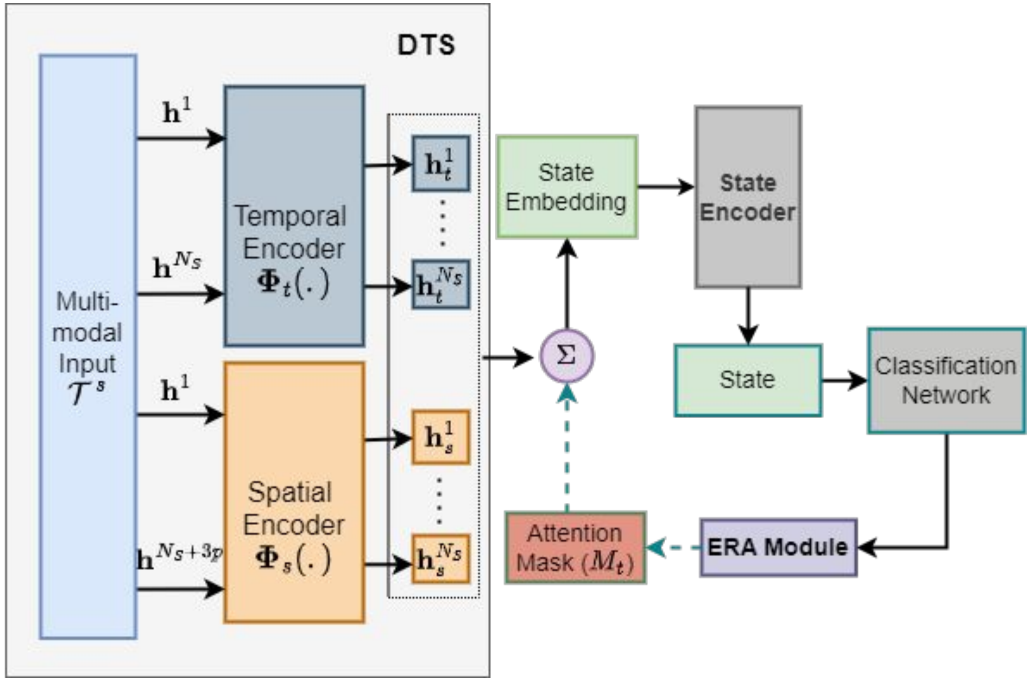
$g_n : n_{th}\ gaze\ instance$ $\quad M_g : Gaze\ feature\ space$ $\quad N_e :\ Length\ of\ sequence\ set\ S$

$t_n : n_{th}\ touch\ instance$ $\quad M_t : Touch\ feature\ space$

Overview of the DTS-ERA Model



(a) Overview of DTS-ERA



(b) DTS structure

Multi-modal Inputs:

$$h^i \; i \in [1, N_s]$$

DTS embeddings:

$$\Phi_{g/t}(h^i) \; i \in [1, N_s]$$

DTS Embeddings Aggregation:

$$d = \mathrm{DTS}([h^1, \cdots, h^{N_s}]) = \frac{1}{N_s} \sum_{i=1}^{N_s} concat[\Phi_g(h^i), \Phi_t(h^i)]$$

RL Attentive DTS Embeddings Aggregation:

$$d_{attn}^t = \frac{1}{N_a} \sum_{k=1}^{N_a} concat[\Phi_g(h^{idx_t^k}), \Phi_t(h^{idx_t^k})]$$

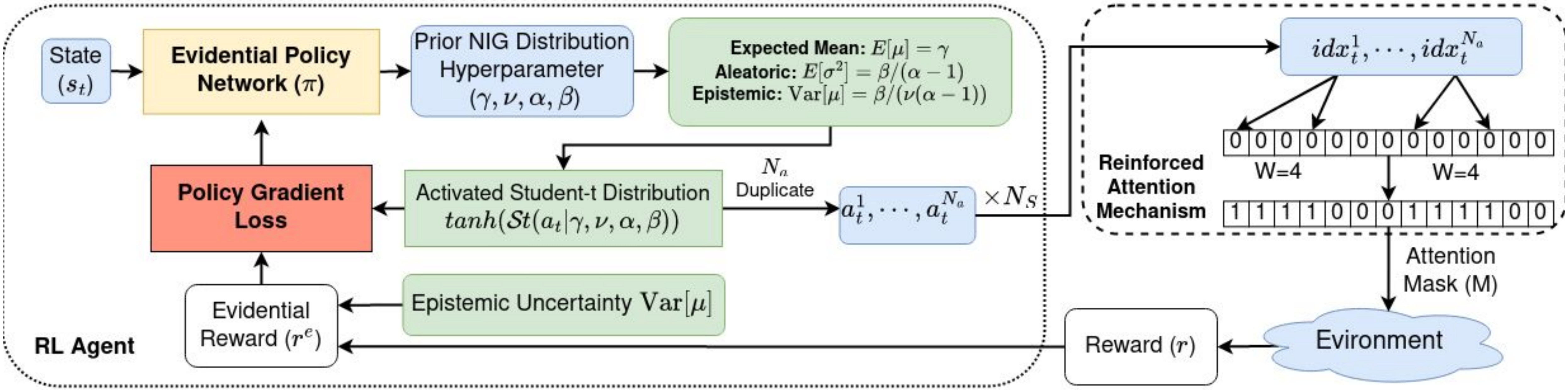
State Embedding:

$$\mathbf{e}_t = \mathrm{concat}(\mathbf{d}, \mathbf{d}_{attn}^t)$$

State:

$$\mathbf{s}_t = \mathrm{SE}(\mathbf{e}_t, \mathbf{s}_{t-1}; \theta_{se})$$

Evidential Reinforced Attention



Action generation: Sampling from a Student-t Distribution (marginal distribution over Gaussian parameters)

$$
\begin{aligned}
& p(\mathbf{a}_t|\gamma, \nu, \alpha, \beta) \\
& = \int_{\sigma^2} \int_{\mu} p(\mathbf{a}_t|\mu, \sigma^2) p(\mu, \sigma^2|\gamma, \nu, \alpha, \beta) \mathrm{d}\mu \mathrm{d}\sigma^2 \\
& = \mathrm{St}\left(\mathbf{a}_t; \gamma, \beta(1+\nu)/(\nu\alpha), 2\alpha\right)
\end{aligned}
$$

Evidential Reward: Combine classification accuracy and epistemic uncertainty to balance exploitation and exploration

$$
\begin{aligned}
r^e(\mathbf{s}_t, \mathbf{a}_t) &= r(\mathbf{s}_t, \mathbf{a}_t) + \lambda \mathtt{epistemic}(\pi_{\theta_e}(\cdot|\mathbf{s}_t)) \\
r(\mathbf{s}_t, \mathbf{a}_t) &= \mathbb{1}\{p_T = y_s\}
\end{aligned}
$$

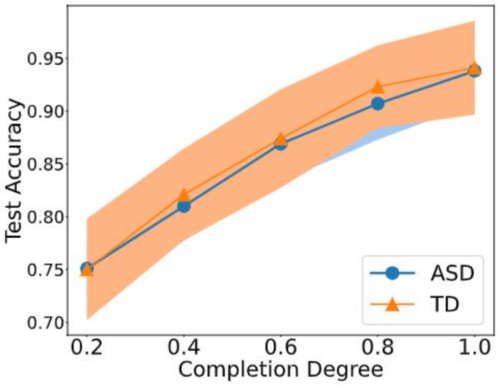
Action designated attended DTS embeddings indexes: $idx_t^k = \sigma(a_t^k) \cdot (N_s - W)$

Experimental Results

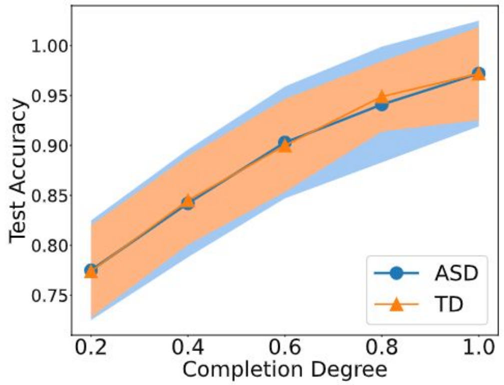
Compare with SOTA:

Dataset	Maze-2D	Maze-3D	Maze-Mixed	Coloring	Word Scanning
TCN	73.1±4.5	64.0±7.1	60.5±5.7	64.0±4.1	68.0±3.7
GRU-FCN	92.3±3.2	94.5±4.1	91.8±4.8	88.0±3.9	91.0±3.8
ResCNN	85.0±5.2	76.0±4.5	82.0±6.2	85.0±4.3	88.0±4.2
InceptionTime	80.1±6.7	72.5±1.6	78.8±4.8	82.0±4.1	77.0±4.2
MiniRocket	70.7±7.1	55.3±3.5	56.3±3.8	65.0±4.8	65.0±5.1
DTS	93.1±3.6	94.6±5.8	92.7±5.6	89.0±4.5	92.5±4.4
DTS-ERA	**94.0±3.8**	**95.3±5.3**	**95.0±5.2**	**91.0±3.9**	**93.5±3.8**

Partial Trajectory Prediction:

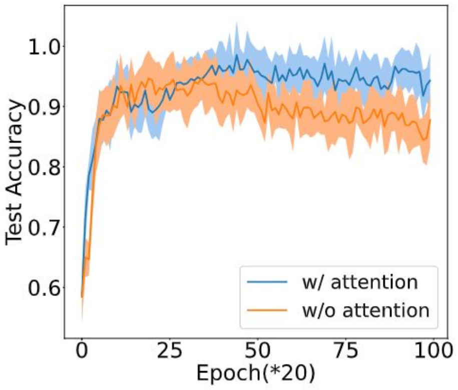


(a) Incomplete data [2D]

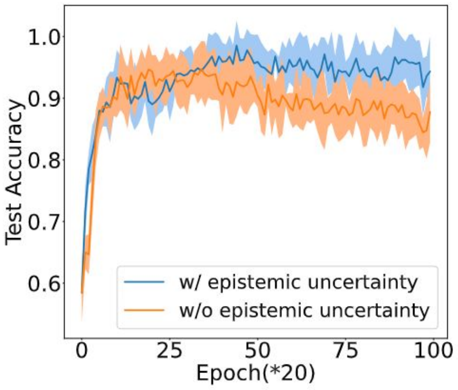


(b) Incomplete data [3D]

Ablation Study:



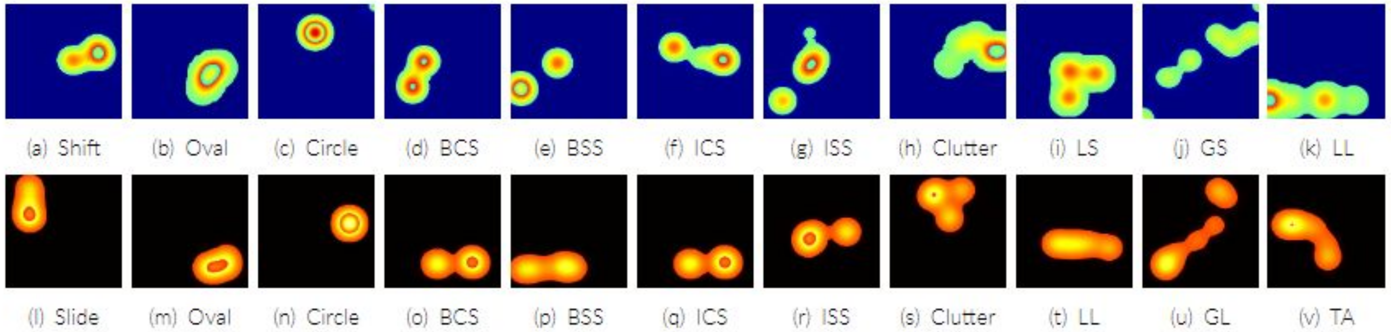
(a) Attention



(b) Evidential exploration

Visualization of Discovered Behavioral Patterns

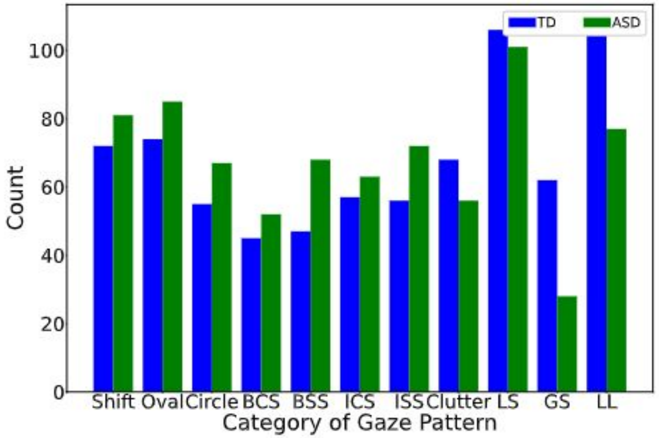
RL Discovered Signature Behavior Patterns (SBPs):

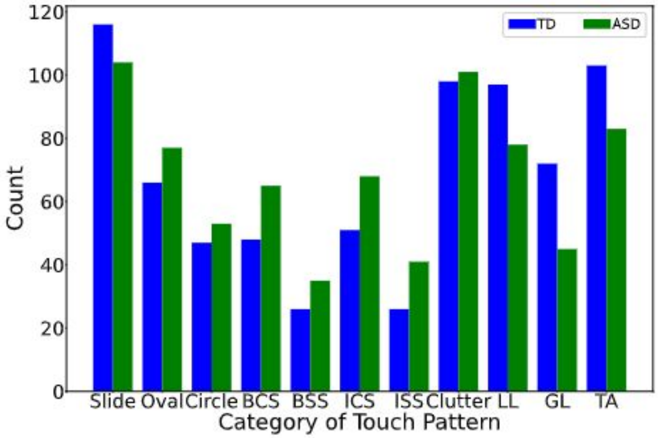


Representative SBPs

Sensory	Shape	Pattern Category	P-value
Gaze	**Concentrated**	Shift	0.0985
		Oval	0.1042
		Circle	0.0965
	Switching	**Balanced Clutter Switching** (BCS)	0.0675
		Balanced Scatter Switching (BSS)	0.0876
		Imbalanced Clutter Switching (ICS)	0.0752
		Imbalanced Scatter Switching (ISS)	0.0923
	Local	**Clutter**	0.0623
		Local Scan (LS)	0.0518
	Global	**Global Scan** (GS)	0.0432
	Line	Local Line (LL)	0.0447
Touch	**Concentrated**	Slide	0.0802
		Oval	0.0857
		Circle	0.0968
	Switching	Balanced Clutter Switching (BCS)	0.0745
		Balanced Scatter Switching (BSS)	0.0736
		Imbalanced Clutter Switching (ICS)	0.0594
		Imbalanced Scatter Switching (ISS)	0.0654
	Local	Clutter	0.0518
		Turning Around (TA)	0.0485
	Line	Local Line (LL)	0.0469
		Global Line (GL)	0.0423

SBP Statistical Analysis:





Welcome to our poster session for more discussion!

THANKS

Paper **Code** **WeChat**

Feel free to contact me via dw7445@rit.edu.