



Attend and Diagnose: Clinical Time Series Analysis Using Attention Models



Motivation

- Lots of data available: electronic health records
- Predict length of hospital stay, mortality etc.
- Attention models: promising results for NLP



Problem description

- State-of-the-art: RNN with LSTM
 - Parallelization is challenging
- Clinical data
 - Irregular sampling
 - Missing data, measurement errors
 - Long range dependencies

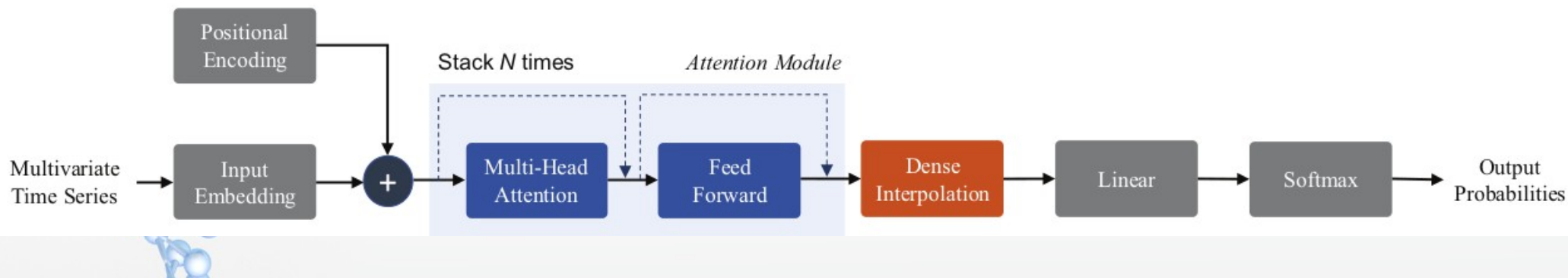


Related work

- Linear Dynamical System
- Gaussian Process
- RNN with LSTM

Method

- SAnD architecture (Simply Attend and Diagnose)
 - Attention mechanism: Masked Self-Attention
 - Positional encoding & Dense interpolation
- SAnD-Multi





Experimental setting

- MIMIC-III Benchmark Dataset

Table 1: Task-specific sample sizes of MIMIC-III dataset.

Benchmark	Train	Validation	Test
Mortality	14,659	3,244	3,236
Decompensation	2,396,001	512,413	523,208
Length of Stay	2,392,950	532,484	525,912
Phenotyping	29,152	6,469	6,281



Experimental results

Table 2: Performance Comparison for the MIMIC-III benchmark tasks, using both single-task and multi-task strategies.

Metrics	Method				
	LR	LSTM	<i>SAnD</i>	LSTM-Multi	<i>SAnD</i> -Multi
Task 1: Phenotyping					
Micro AUC	0.801	0.821	0.816	0.817	0.819
Macro AUC	0.741	0.77	0.766	0.766	0.771
Weighted AUC	0.732	0.757	0.754	0.753	0.759
Task 2: In Hospital Mortality					
AUROC	0.845	0.854	0.857	0.863	0.859
AUPRC	0.472	0.516	0.518	0.517	0.519
min(Se, P+)	0.469	0.491	0.5	0.499	0.504
Task 3: Decompensation					
AUROC	0.87	0.895	0.895	0.900	0.908
AUPRC	0.2132	0.298	0.316	0.319	0.327
min(Se, P+)	0.269	0.344	0.354	0.348	0.358
Task 4: Length of Stay					
Kappa	0.402	0.427	0.429	0.426	0.429
MSE	63385	42165	40373	42131	39918
MAPE	573.5	235.9	167.3	188.5	157.8



Conclusion

- Computationally superior
- Comparable results/superior