Prognosis

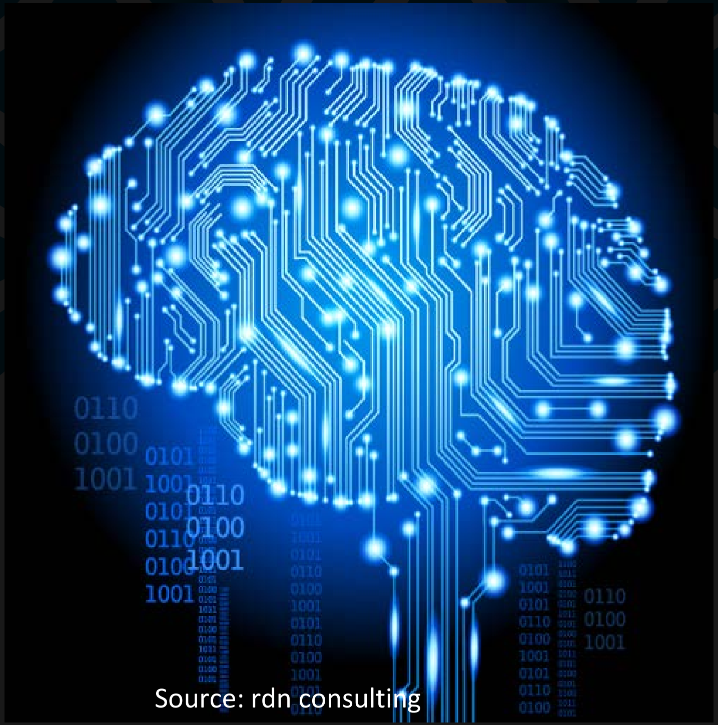
Care

Deep learning for episodic interventional data

Truyen Tran
Deakin University

truyen.tran@deakin.edu.au
truyentran.github.io
@truyenoz
letdataspeak.blogspot.com
goo.gl/3jJ1O0

Melbourne, Aug 2018



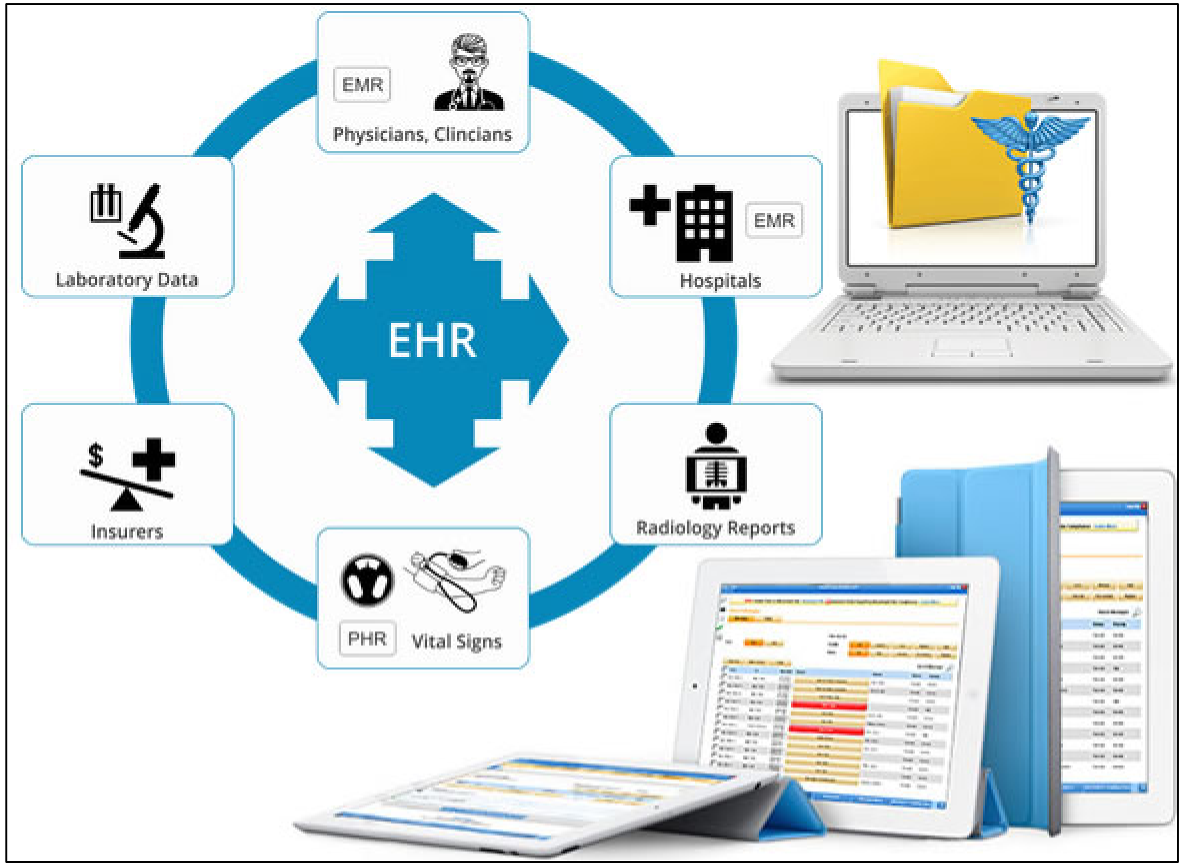
Source: rdn consulting

Setting: Three interacting processes in EMR

Disease progression

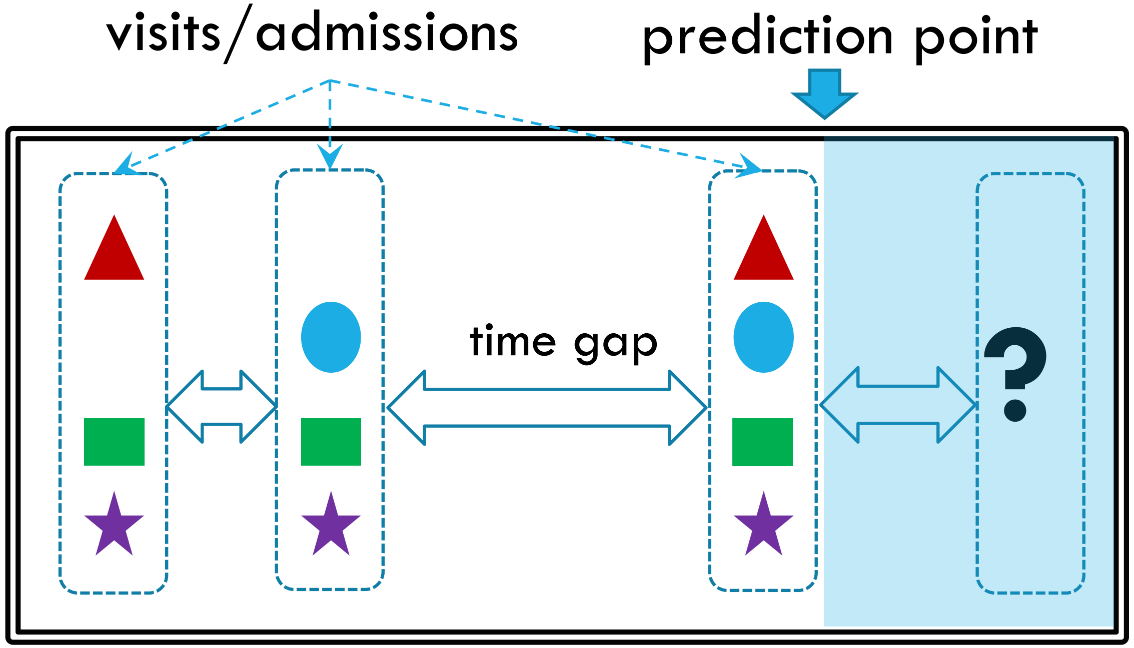
Interventions & care processes

Recording rules



Source: medicalbillingcodings.org

Challenge: Complexity



Long-term dependencies

Irregular timing

Mixture of discrete codes and continuous measures

Complex interaction of diseases and care processes

Rich domain knowledge & ontologies

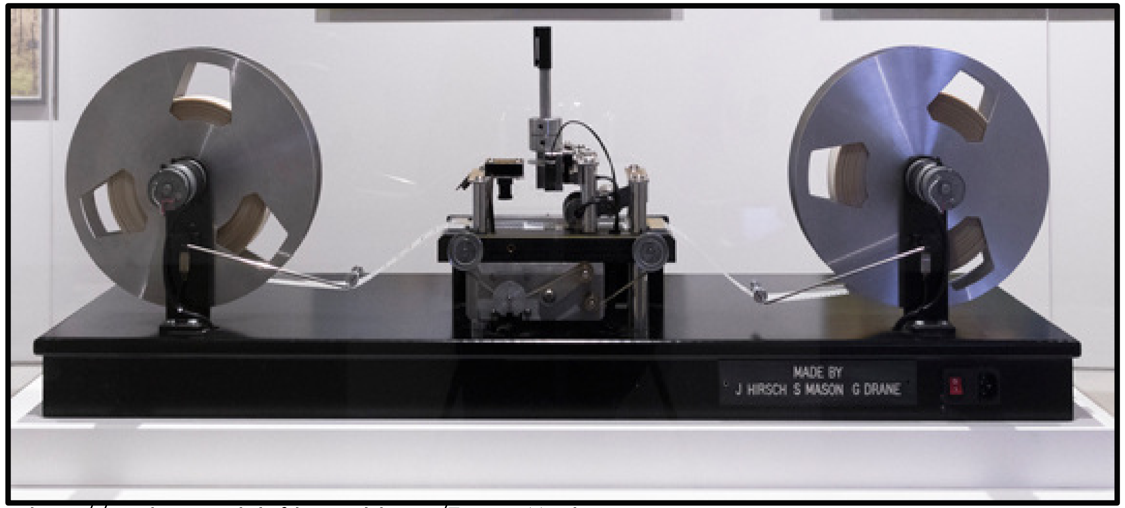
Data are not created equally important

Models must be accurate AND explainable

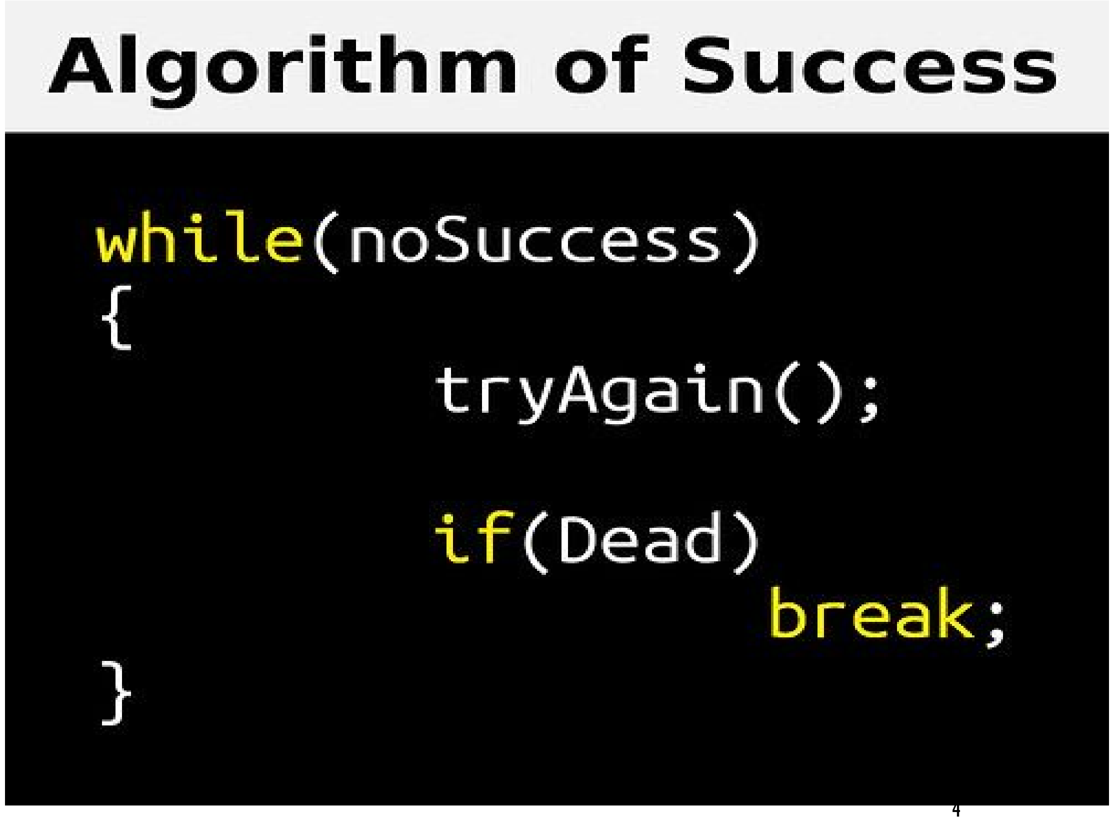
Hypothesis: Healthcare is Turing computational

Healthcare processes as executable computer program obeying hidden "grammars"

The "grammars" are learnable through observational data



http://workingmodeloftheworld.com/Turing-Machine



https://www.pinterest.com.au/pin/140738175619879 32/

Model: Trainable algebraic system

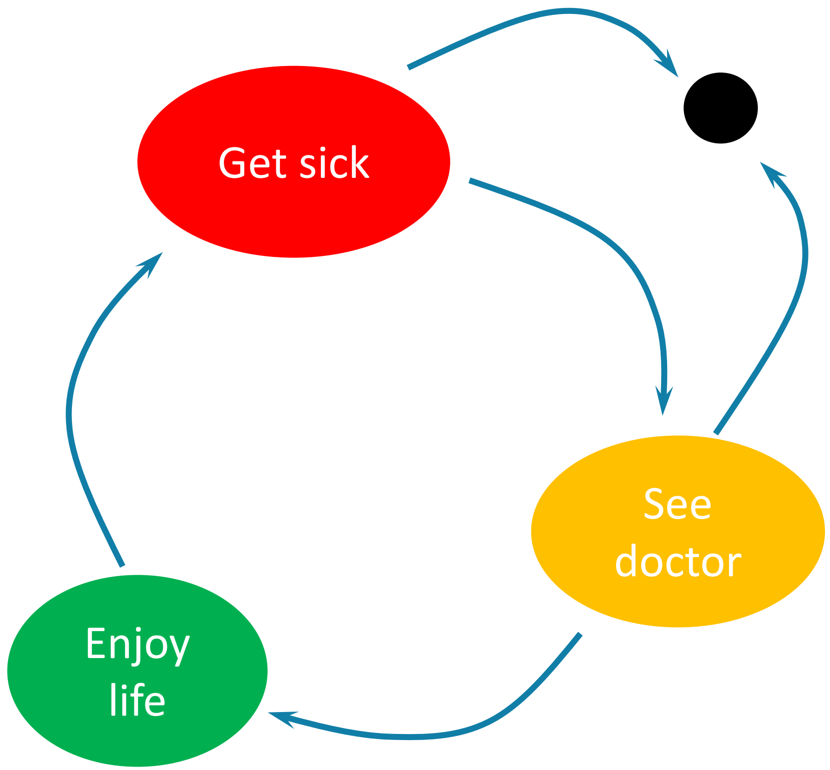
Health dynamics as a system of transitions of forgettable illness states

Treatments "shift" illness state from one point to another

Medical entities (e.g., disease, treatment, doctor) **as algebraic objects** (e.g., vector, matrix and function)

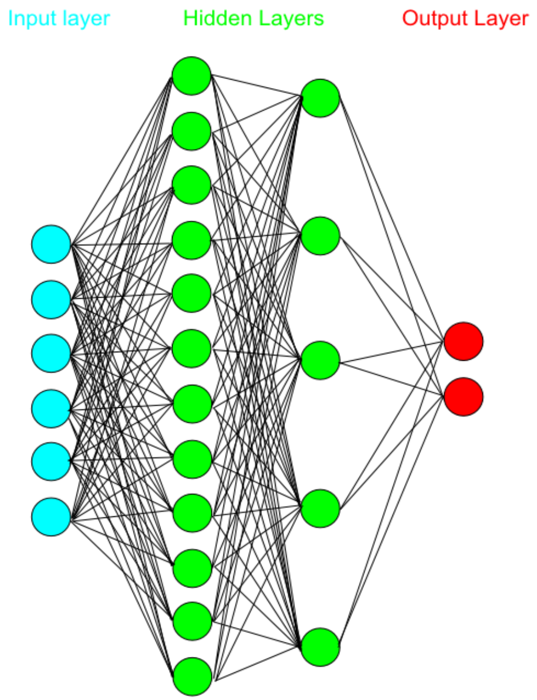
Importance of historical events are person-specific

Training by minimizing prediction loss



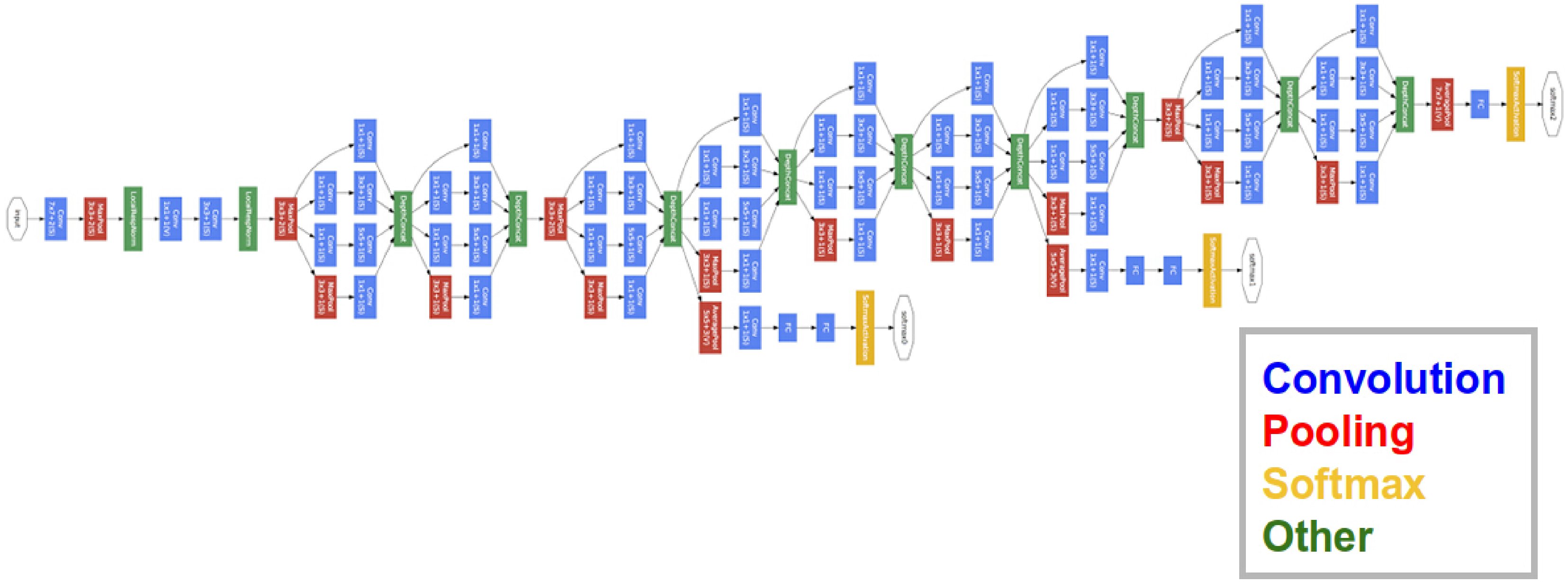
Realisation: Deep learning

1986

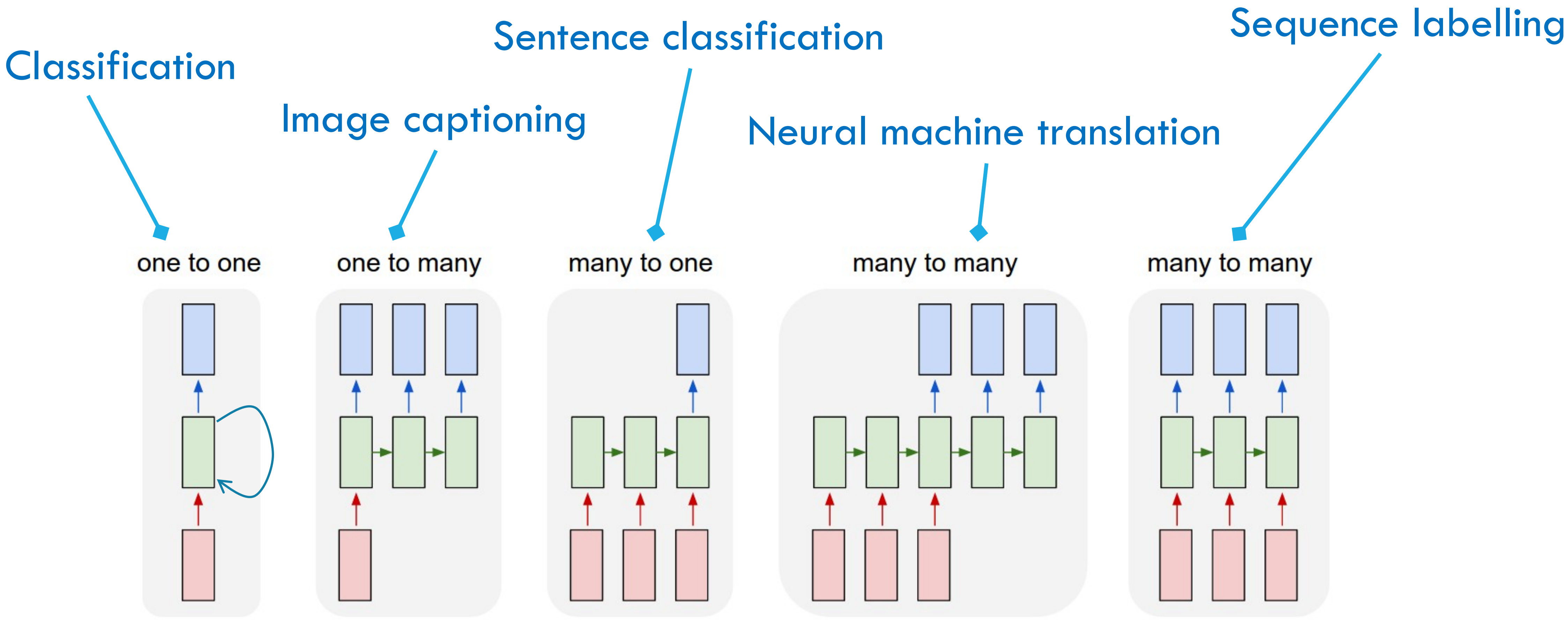


http://blog.refu.co/wp-content/uploads/2009/05/mlp.png

2016

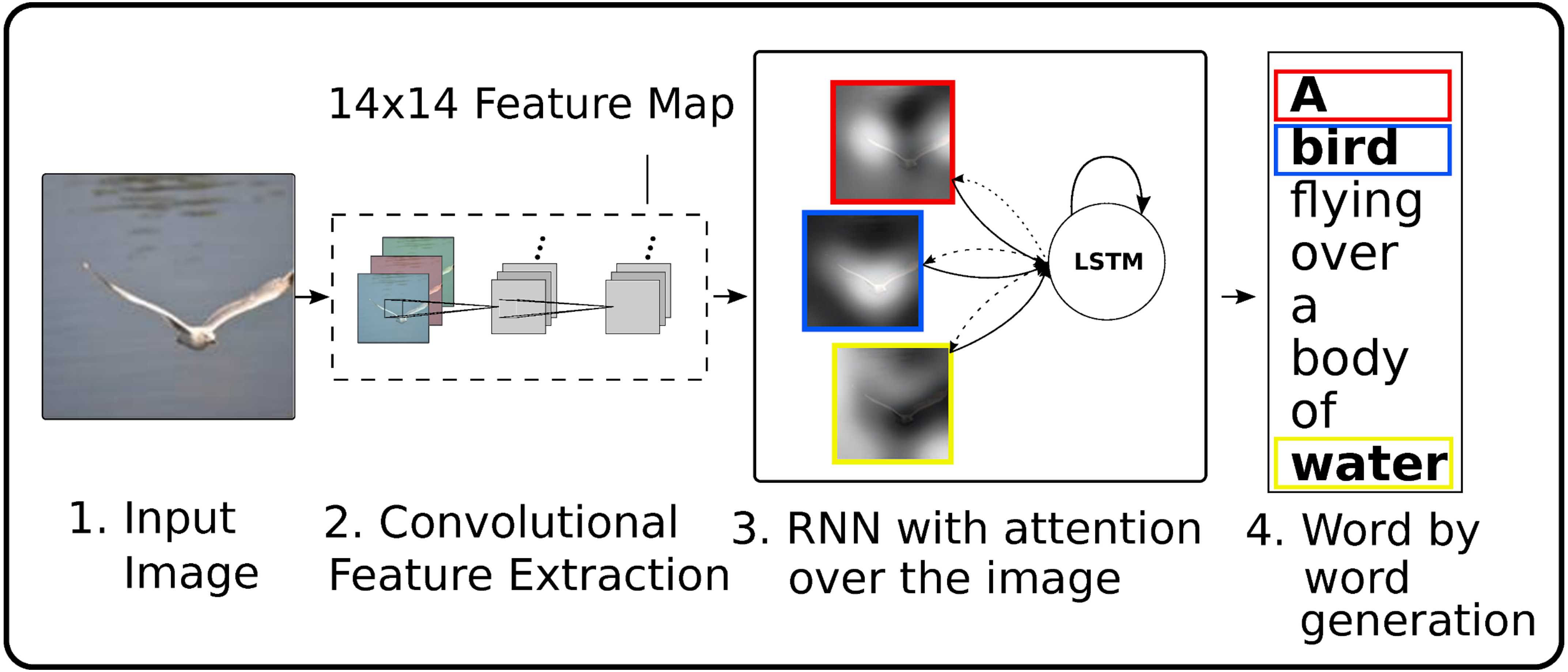


Recurrent neural networks



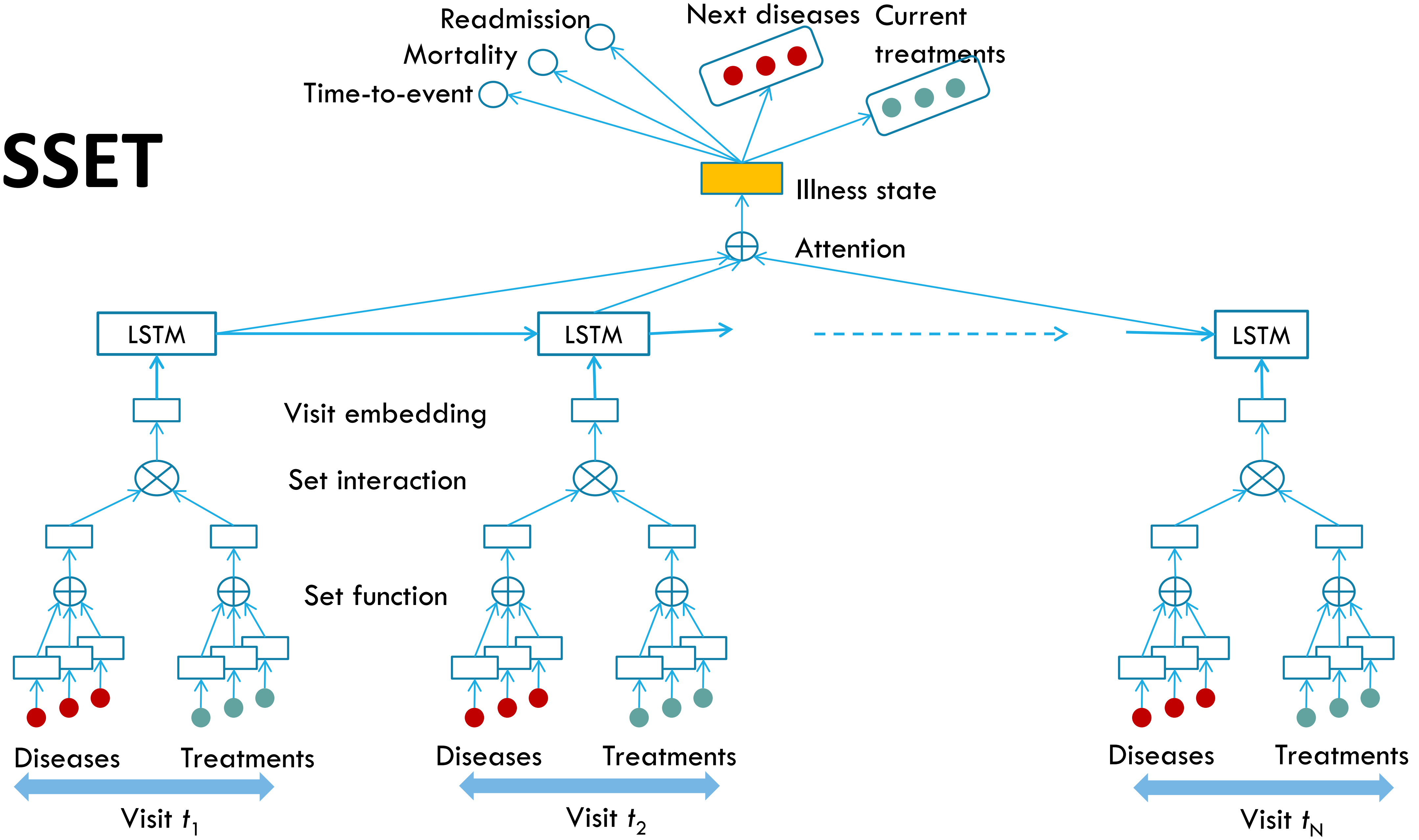
Source: http://karpathy.github.io/assets/rnn/diags.jpeg

Attention mechanism



Show, Attend and Tell: Neural Image Caption Generation with Visual Attention, K. Xu , J. Ba, R. Kiros, K. Cho, A. Courville, R. Salakhutdinov, R. Zemel, Y. Bengio

RESSET



#REF: Phuoc Nguyen, Truyen Tran, and Svetha Venkatesh. "Resset: A Recurrent Model for Sequence of Sets with Applications to Electronic Medical Records." *IJCNN* (2018).

Health = *Recurrence*(Set(Illness) – Set(Intervention))

health

non-linearity

illness

intervention

$$\underbrace{g}_{\text{health}}(\boldsymbol{d}_t, \boldsymbol{p}_t) = \underbrace{\rho}_{\text{non-linearity}}(\Delta) \quad \text{where} \quad \Delta = \underbrace{\boldsymbol{d}_t}_{\text{illness}} - \underbrace{\boldsymbol{p}_t}_{\text{intervention}}$$

$$f(S) \leftarrow \frac{\bar{\boldsymbol{e}}_S}{\epsilon + \|\bar{\boldsymbol{e}}_S\|} \quad \text{where} \quad \bar{\boldsymbol{e}}_S = \max\left(\boldsymbol{0}, \sum_{i \in S} \boldsymbol{e}_i\right)$$

set function ~ permutation invariant

Data: Barwon Health, Geelong Australia (2002-2013)

Statistics	Diabetes	Mental health
# patients	7,191	6,109
# visits	53,208	52,049
% male	55.5	49.4
median age	73	37
# diseases	243	247
# treatments	1,118	1,071

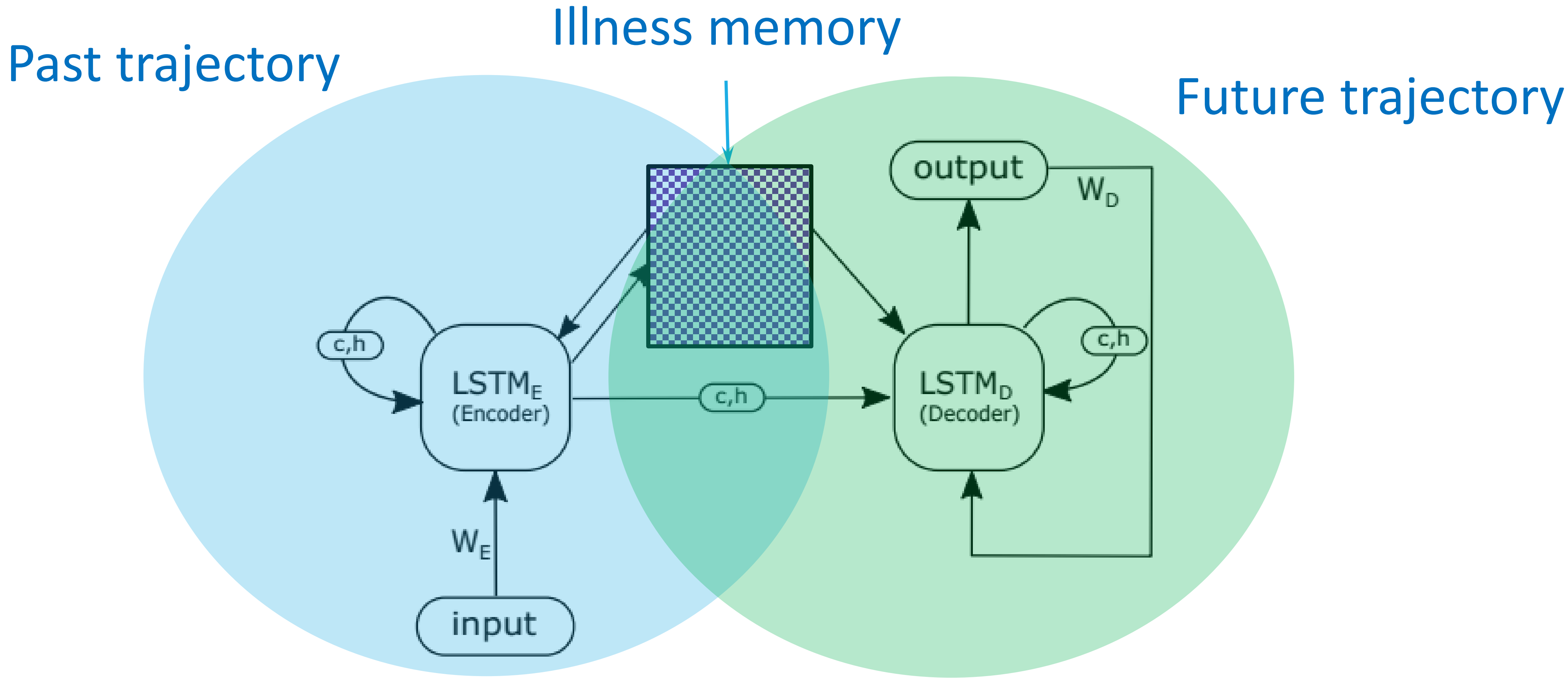
Results: Treatment recommendation

Method	Diabetes			Mental health		
	P@1	P@2	P@3	P@1	P@2	P@3
BOW+LR	0.608	0.481	0.419	0.516	0.4382	0.395
Deepr	0.634	0.463	0.395	0.615	0.532	0.466
LSTM	0.694	0.535	0.446	0.614	0.507	0.427
Resset						
– Implicit interaction	**0.738**	**0.564**	**0.492**	0.692	0.582	**0.498**
– Additive interaction	**0.74**	**0.567**	**0.486**	**0.708**	**0.588**	0.496
– Subtractive interaction	0.704	0.553	0.48	**0.7**	**0.591**	**0.51**
– Multiplicative interaction	0.65	0.484	0.401	0.553	0.511	0.428
– Add. interaction with exp smoothing	0.726	**0.564**	0.465	0.654	0.537	0.458
– Sub. interaction with exp smoothing	0.730	0.561	0.465	0.641	0.528	0.452

Results: Disease progression

Method	Diabetes			Mental health		
	P@1	P@2	P@3	P@1	P@2	P@3
BOW+LR	0.508	0.441	0.393	0.396	0.350	0.323
Deepr	0.496	0.42	0.397	0.424	0.392	0.346
LSTM	0.541	0.476	0.417	0.466	0.430	0.372
`Resset`						
– Implicit interaction	0.530	0.478	0.438	**0.504**	**0.471**	**0.406**
– Additive interaction	0.528	0.496	0.449	0.488	0.448	0.392
– Subtractive interaction	0.533	0.491	0.444	**0.494**	**0.469**	**0.41**
– Multiplicative interaction	0.496	0.44	0.401	0.453	0.406	0.362
– Add. interaction with exponential smoothing	**0.563**	**0.513**	**0.459**	0.468	0.429	0.373
– Sub. interaction with exponential smoothing	**0.567**	**0.516**	**0.46**	0.47	0.43	0.376

Towards a differentiable Turing machine for health



#REF: Hung Le, Truyen Tran, and Svetha Venkatesh. "Dual Control Memory Augmented Neural Networks for Treatment Recommendations", PAKDD18.

Results: MIMIC-III data

Model	Procedure Output		Drug Output	
	Precision	Jaccard	Precision	Jaccard
Logistic Regression	0.256	0.185	0.412	0.311
Random Forest	0.276	0.199	0.491	0.405
Seq2Seq	0.263	0.196	0.220	0.138
Seq2Seq with attention	0.272	0.204	0.224	0.142
DNC	0.285	0.214	0.577	0.529
DCw-MANN	**0.292**	**0.221**	**0.598**	**0.556**

Wrapping up

Healthcare can be viewed as an algebraic system, which can be realized as a learnable Turing program

Dynamic diseases-treatments interaction demands new models

Deep learning is a viable solution

Question: where is the likelihood function?

2012

The Team @ Deakin

DEAKIN
UNIVERSITY AUSTRALIA
Worldly

Epworth
HealthCare

Barwon
Health