Discovery

Diagnosis

Prognosis

Care

Deep learning for biomedicine II

Truyen Tran
Deakin University

truyen.tran@deakin.edu.au

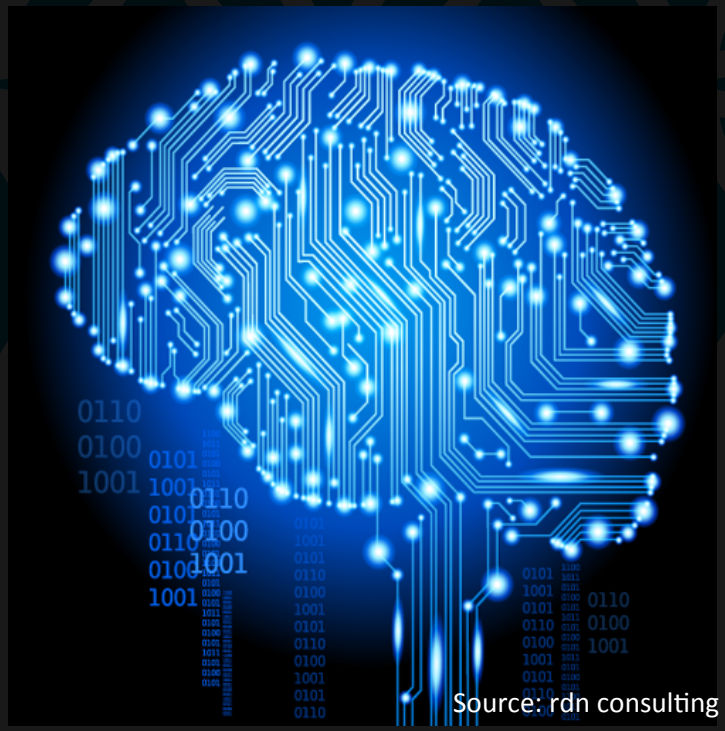
truyentran.github.io

@truyenoz

letdataspeak.blogspot.com

goo.gl/3jJ1O0

Seoul, Nov 2017



Resources

Slides and references:

- https://truyentran.github.io/acml17-tute.html

Key survey paper (updated frequently):

- **Ching, Travers, et al. "Opportunities And Obstacles For Deep Learning In Biology And Medicine." *bioRxiv* (2017): 142760**

Agenda

Topic 1: Introduction (20 mins)

Topic 2: Brief review of deep learning (25 mins)

- Classic architectures
- Capsules
- Graphs
- Memory-augmented nets

Topic 3: Genomics (25 mins)

- Nanopore sequencing
- Genomics modelling

QA (10 mins)

Break (20 mins)

Topic 4: Biomedical imaging (15 mins)

- Cellular imaging
- Diagnostics imaging
- EEG/ECG

Topic 5: Healthcare (25 mins)

- Time-series of physio measures
- Trajectories prediction

Topic 6: Generative biomed (30 mins)

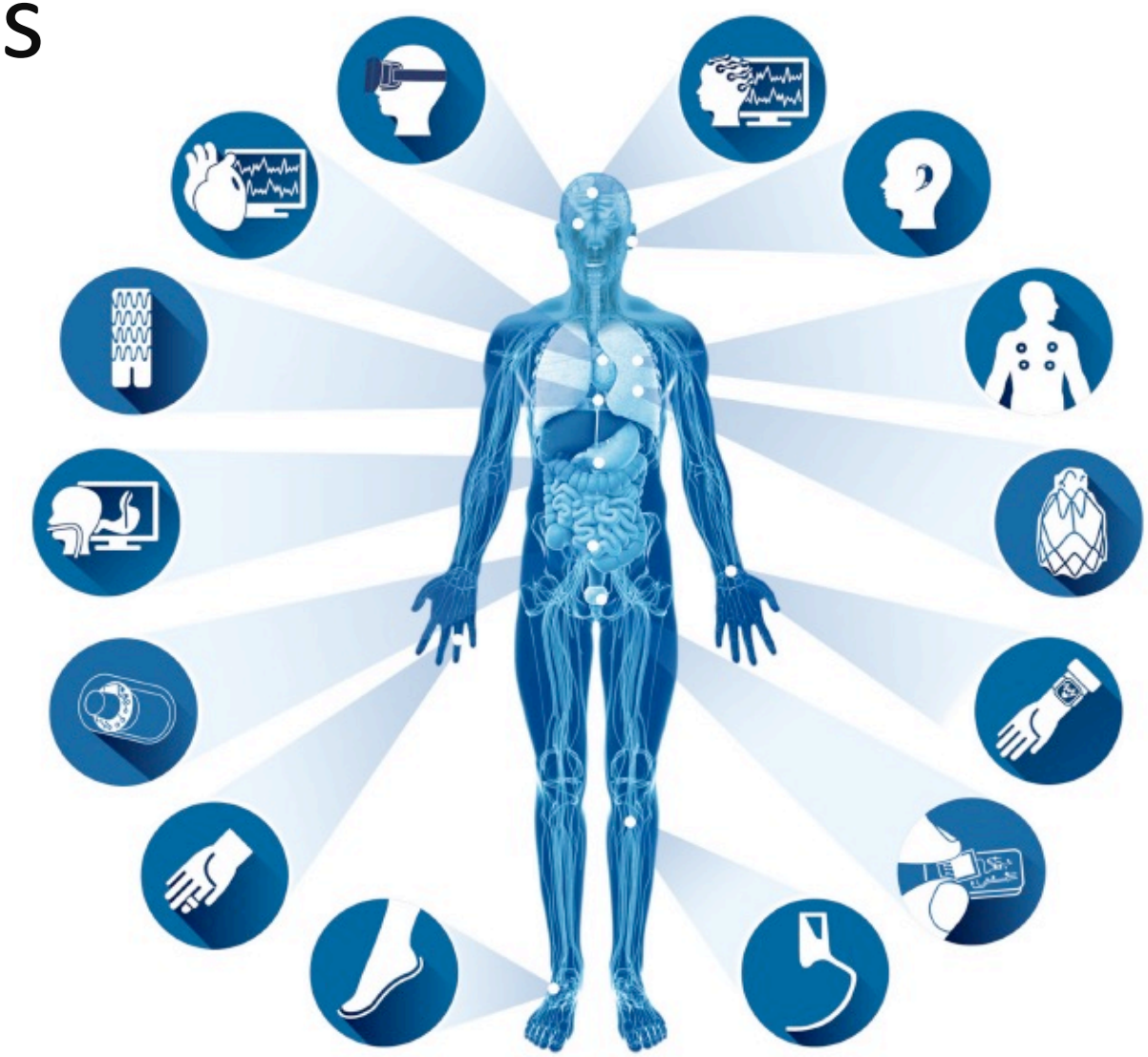
- Few-shot learning
- Generative models
- Drug design
- Future outlook

QA (10 mins)

Sensing technologies and data

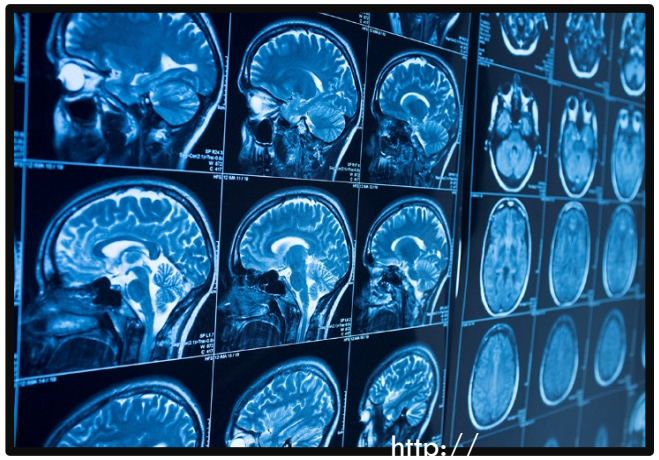
Raw signals are ideal candidates for deep learning

Speech & vision techniques can be applied with minimal changes



#REF: Ravì, Daniele, et al. "Deep learning for health informatics." *IEEE journal of biomedical and health informatics* 21.1 (2017): 4-21.

The state of biomedical imaging



As estimated by IBM, 90% data of healthcare is imaging.

Biomedical imaging is perhaps the most ready area for current DL techniques:

- **It typically means CNN!**
- **The game is in the data acquisition and problem definition/transformation**

Examples of application areas:

- Cellular imaging
- Tumor Detection & tracking
- Blood Flow Quantification and Visualization
- Medical Interpretation
- Diabetic Retinopathy

Tasks

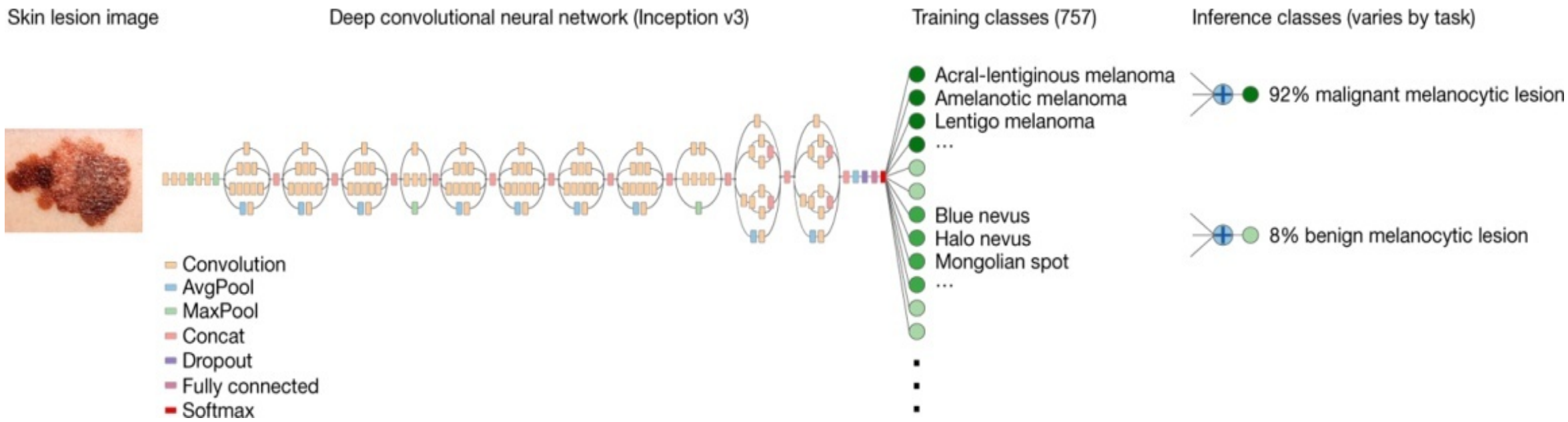
- Classification
- Segmentation
- Localization

Challenges

- Low quality
- Very high resolution
- Tiny localized areas

https://www.healthcare-informatics.com/news-item/analytics/ibm-unveils-watson-powered-imaging-solutions-rsna

Skin cancer diagnosis using **inception**-v3



- 129,450 clinical images
- 2,032 different diseases
- Test against 21 board-certified dermatologists
- Use case 1: keratinocyte carcinomas versus benign seborrheic keratoses;
- Use case 2: malignant melanomas versus benign nevi.

#REF: A Esteva *et al. Nature* 1–4 (2017) doi:10.1038/nature21056

Microscopy + mobile phone + CNN

Highly relevant for:

- Developing countries
- Rural areas

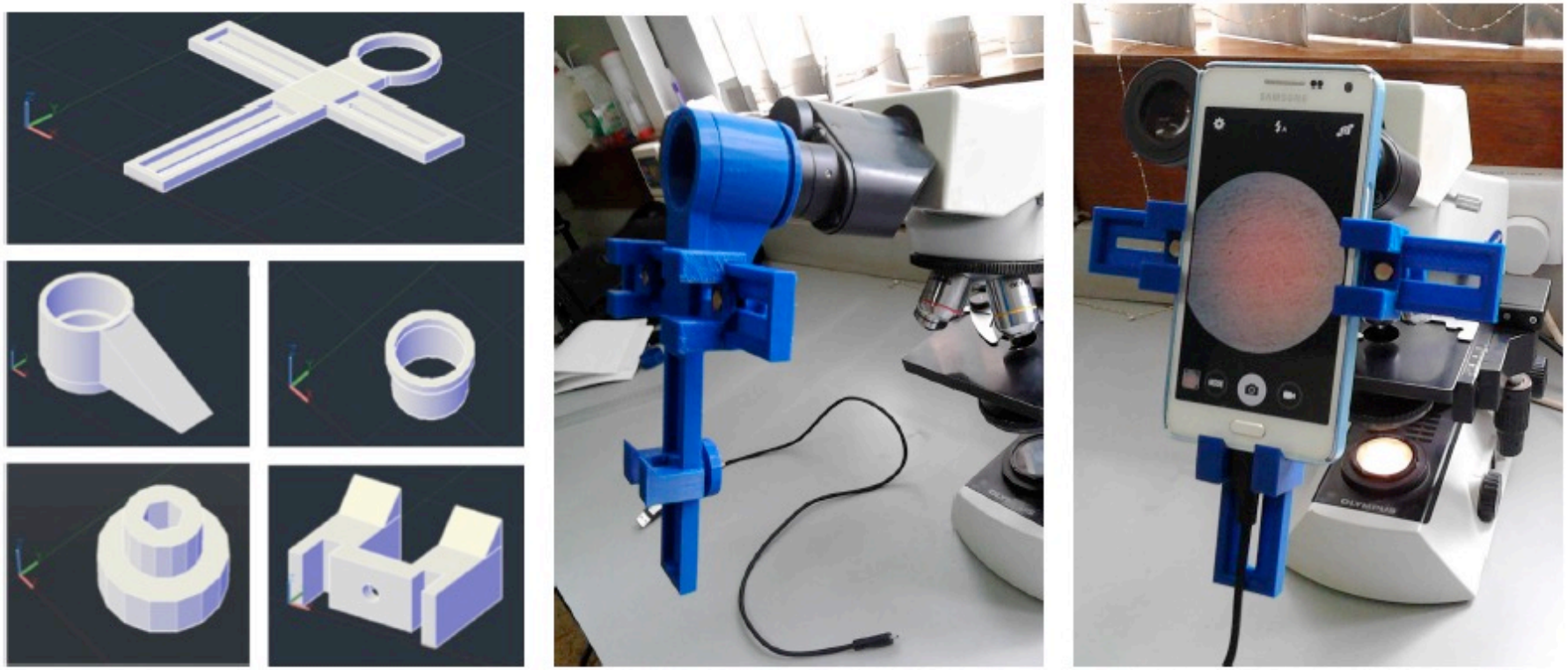


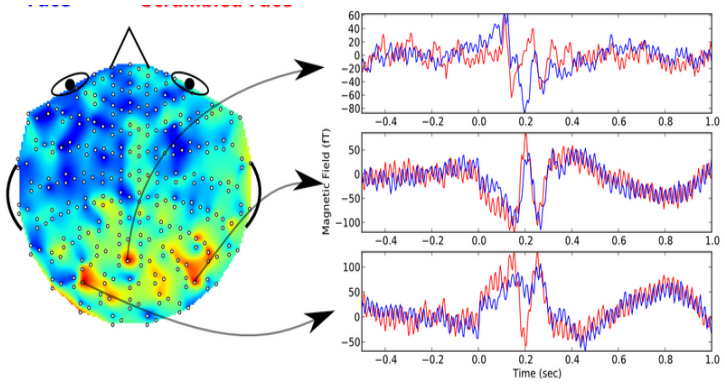
Figure 1: Microscope smartphone adapter: design of components (left), 3D-printed adapter mounted on microscope (center), smartphone inserted into adapter (right).

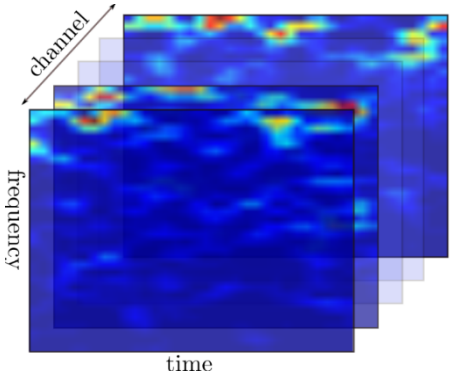
#REF: Quinn, John A., et al. "Deep convolutional neural networks for microscopy-based point of care diagnostics." *Machine Learning for Healthcare Conference*. 2016.

EEG → Tensor RBM for alcoholic diagnosis

control

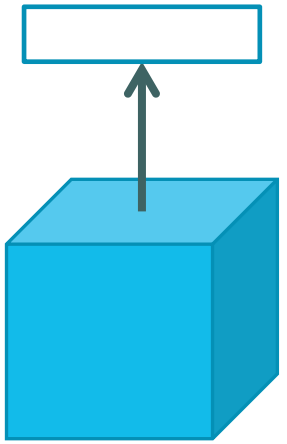
alcoholic





3D Spectrogram

Subject representation



Tensor Restricted Boltzmann Machine

#Ref: Tu D. Nguyen, Truyen Tran, D. Phung, and S. Venkatesh, Tensor-variate Restricted Boltzmann Machines, *AAAI* 2015.

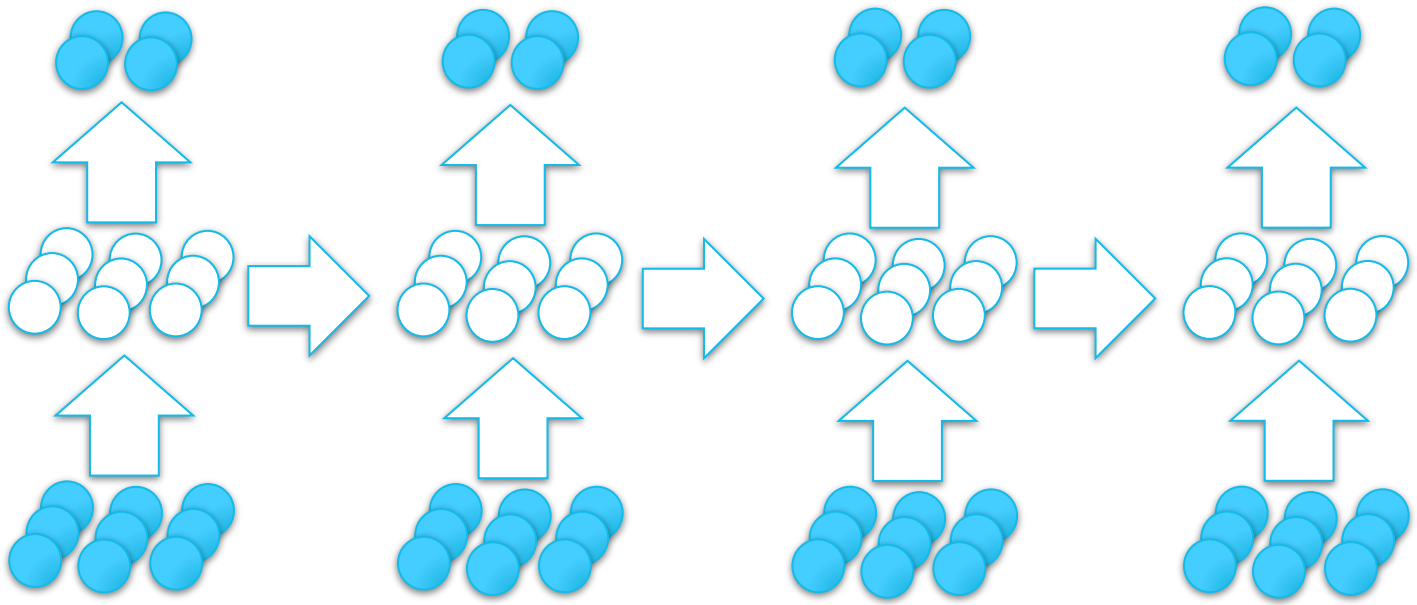
EEG → Matrix LSTM → Classification

EEG segments as matrices

Temporal dynamics as recurrence

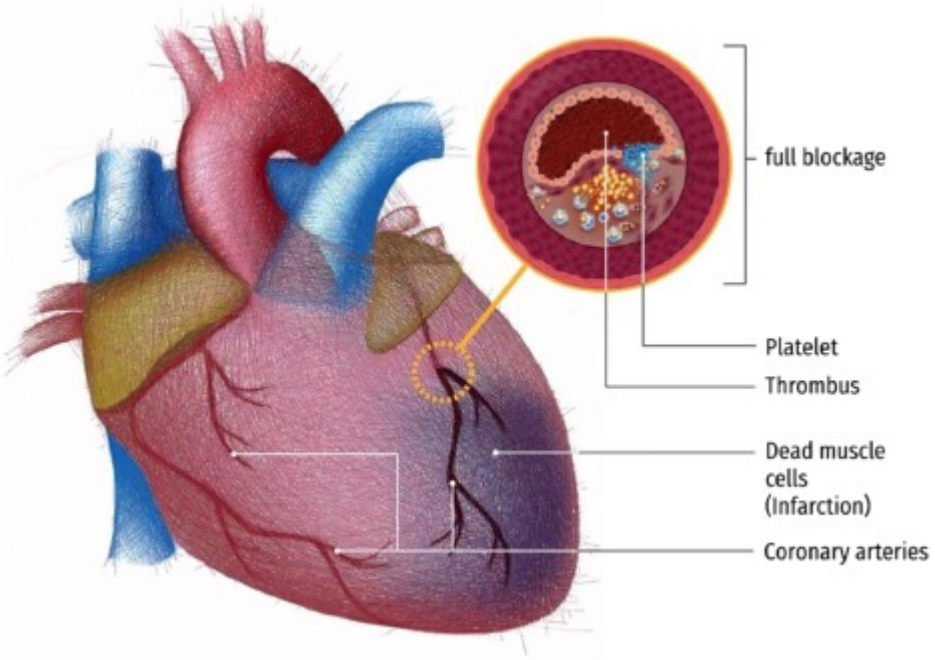
Recurrent dynamics

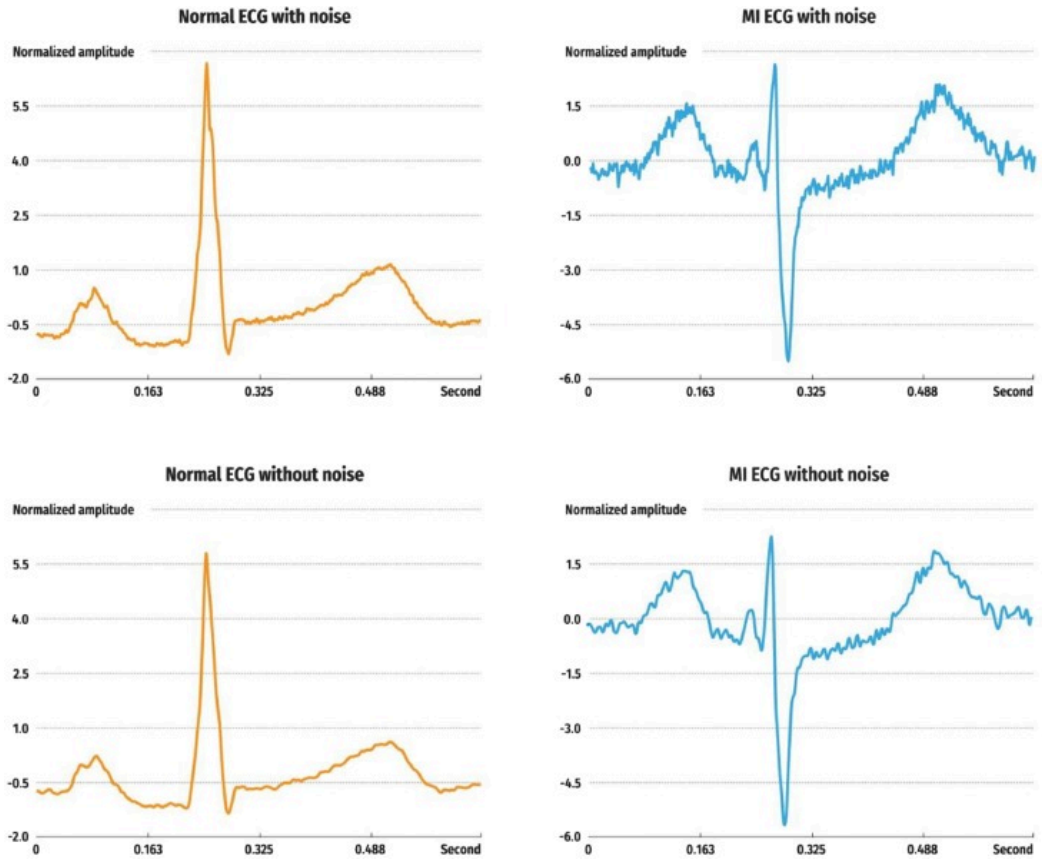
$$H_t = \sigma(U_x^\intercal X_t V_x + U_h^\intercal H_{t-1} V_h + B)$$



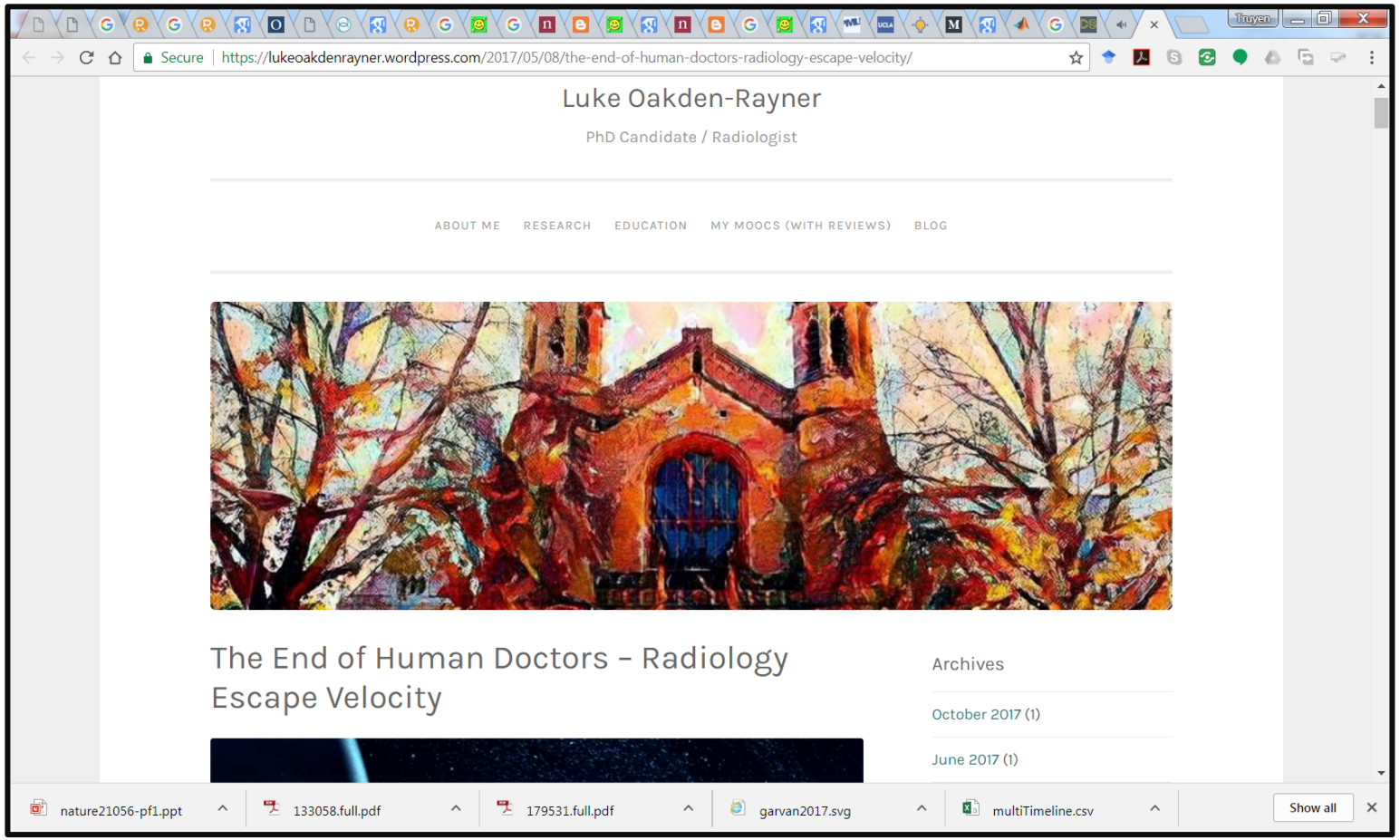
#REF: Kien Do, Truyen Tran, Svetha Venkatesh, "Learning Recurrent Matrix Representation", *Third Representation Learning for Graphs Workshop* (ReLiG 2017)

ECG → CNN for heart attack detection





#REF: Acharya, U. Rajendra, et al. "Application of deep convolutional neural network for automated detection of myocardial infarction using ECG signals." *Information Sciences* 415 (2017): 190-198.



“They should stop training radiologists now.”

Geoff Hinton (as of April 2017)

https://www.newyorker.com/magazine/2017/04/03/ai-versus-md

Agenda

Topic 1: Introduction (20 mins)

Topic 2: Brief review of deep learning (25 mins)

- Classic architectures
- Capsules
- Graphs
- Memory-augmented nets

Topic 3: Genomics (25 mins)

- Nanopore sequencing
- Genomics modelling

QA (10 mins)

Break (20 mins)

Topic 4: Biomedical imaging (15 mins)

- Cellular imaging
- Diagnostics imaging
- EEG/ECG

Topic 5: Healthcare (25 mins)

- Time-series of physio measures
- Trajectories prediction

Topic 6: Generative biomed (30 mins)

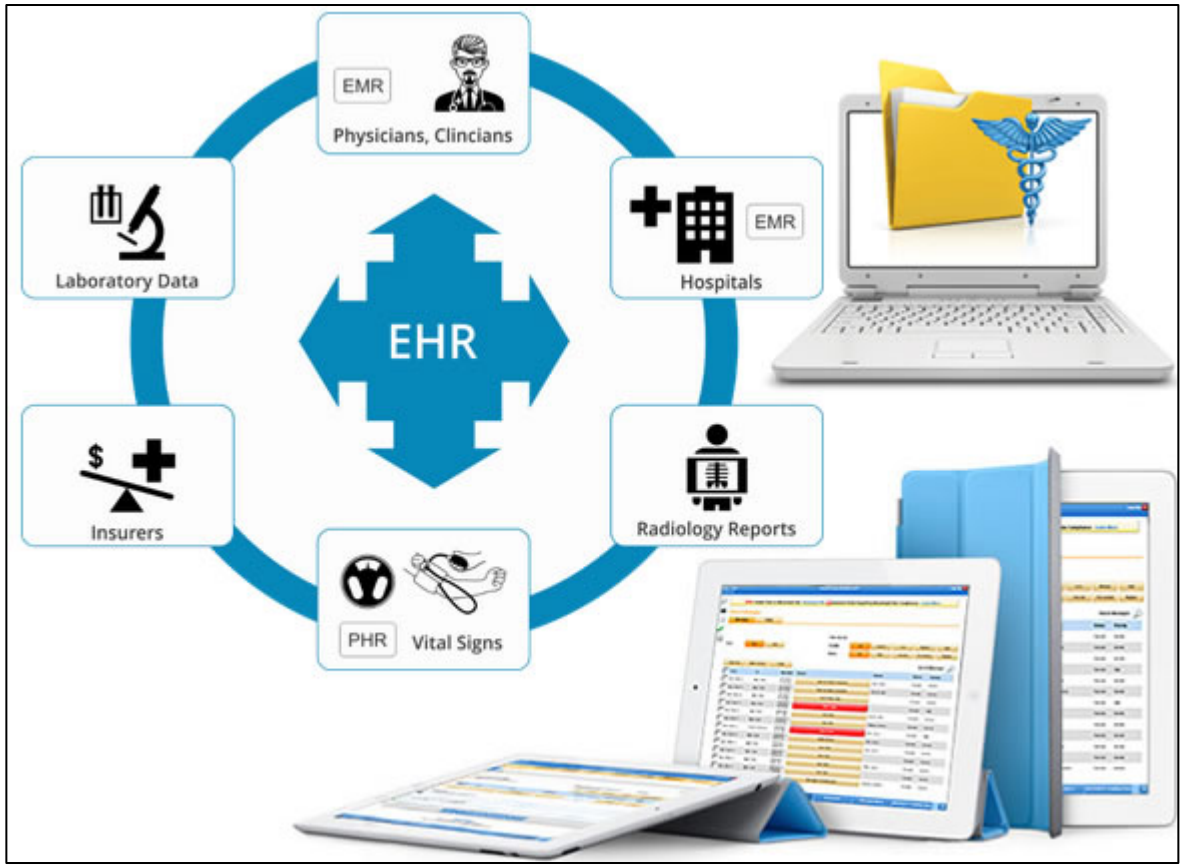
- Few-shot learning
- Generative models
- Drug design
- Future outlook

QA (10 mins)

Modeling electronic medical records (EMR)

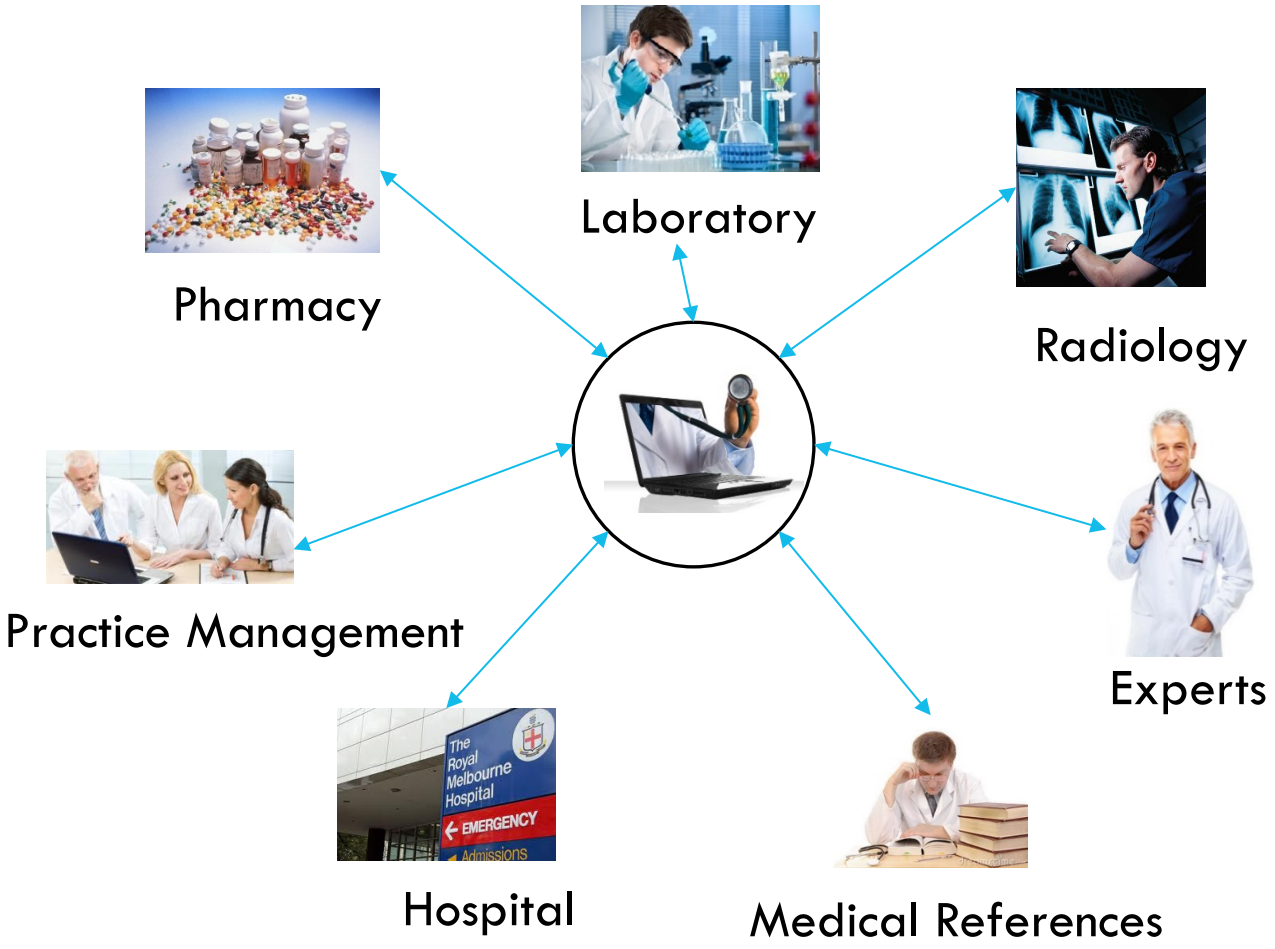
Need to model the healthcare processes, which are interactions of:

- Disease progression
- Interventions & care processes
- Recording processes (Electronic Medical/Health Records)



Source: medicalbillingcodings.org

EMR Connects Services: System of Systems



Five main functions

- Integrated view of patient data
- Clinical decision support
- Clinician order entry
- Access to knowledge resources
- Integrated communication and reporting support

Clinical Decision Supports

Support protocol/planning of treatment/discharge.

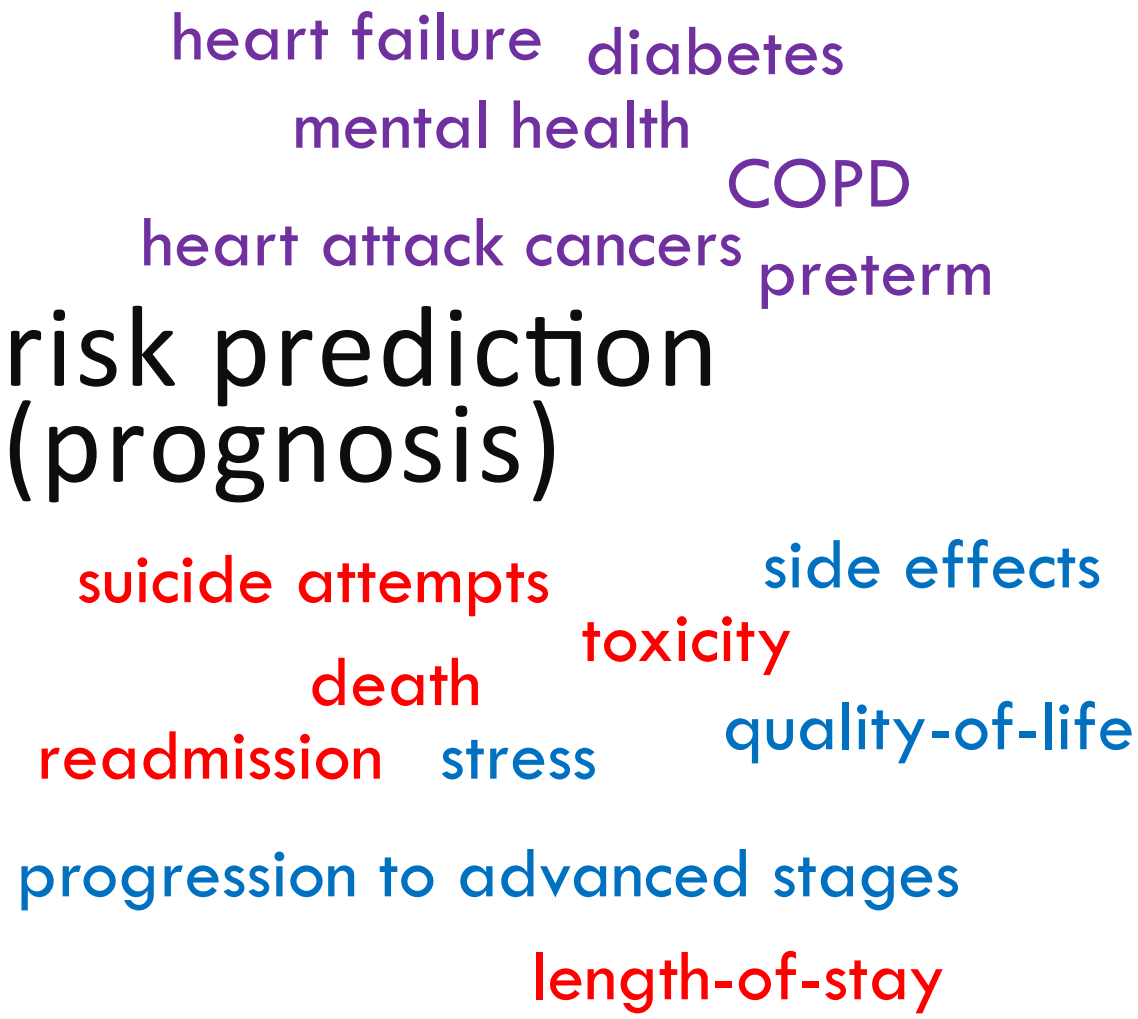
Suggest course of actions:

- E.g., medication/dose/duration.

Estimate risk & predict outcomes.

Alert/reminder.

Support (semi) automated diagnosis.



Warning: leakage!

Make sure the patients are counted AFTER first diagnosis

- Often, we have future data as well
- Retrospective nature

Never use outcomes to do anything, except for training the model

Our early suicide attempt classification from assessments was a form of leakage:

- Any attempt in history is considered as an outcome. BUT:
- Previous attempts were accounted in current assessment already!

Preprocessing: Data normalization & dictionary compression

Drugs & tests

- Drug companies offer different brand names of the essentially the same drug
- DDD/ATC is the central register for the medication classes, maintained by WHO
- Several test names may be the same

It may not be robust to use the original “vocabularies”

- Tens of thousands of ICD-codes, thousands of procedures, hundreds of DRGs, thousands of medication classes
- Codes are usually organized in hierarchy
- Choosing the right hierarchy is statistical issue

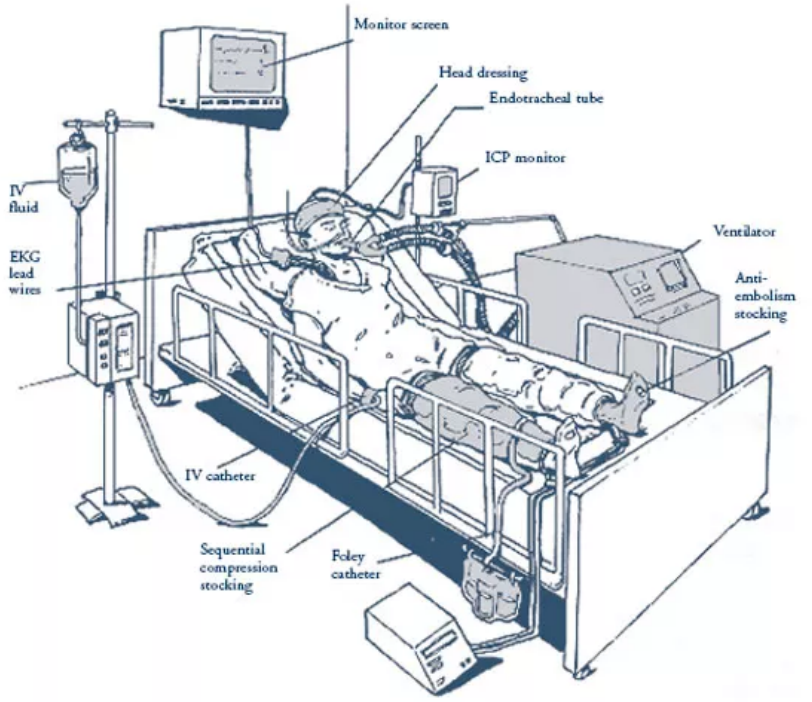
Attend to risks in Intensive Care Unit (ICU)

The needs

- Accuracy
- Interpretability
- As early as possible

The process:

- Irregular time-series → Regular time-steps → Data imputation → Bi-LSTM → **Multiple attentions** → Classification



Source: healthpages.org

Time,Parameter,Value
00:00,RecordID,132539
00:00,Age,54
00:00,Gender,0
00:00,Height,-1
00:00,ICUType,4
00:00,Weight,-1
00:07,GCS,15
00:07,HR,73
00:07,NIDiasABP,65
00:07,NIMAP,92.33
00:07,NISysABP,147
00:07,RespRate,19
00:07,Temp,35.1
00:07,Urine,900
00:37,HR,77
00:37,NIDiasABP,58
00:37,NIMAP,91
00:37,NISysABP,157
00:37,RespRate,19
00:37,Temp,35.6
00:37,Urine,60

Data: **Physionet 2012**

#REF: Phuoc Nguyen, Truyen Tran, Svetha Venkatesh, "Deep Learning to Attend to Risk in ICU", *IJCAI'17 Workshop on Knowledge Discovery in Healthcare II: Towards Learning Healthcare Systems* (KDH 2017).

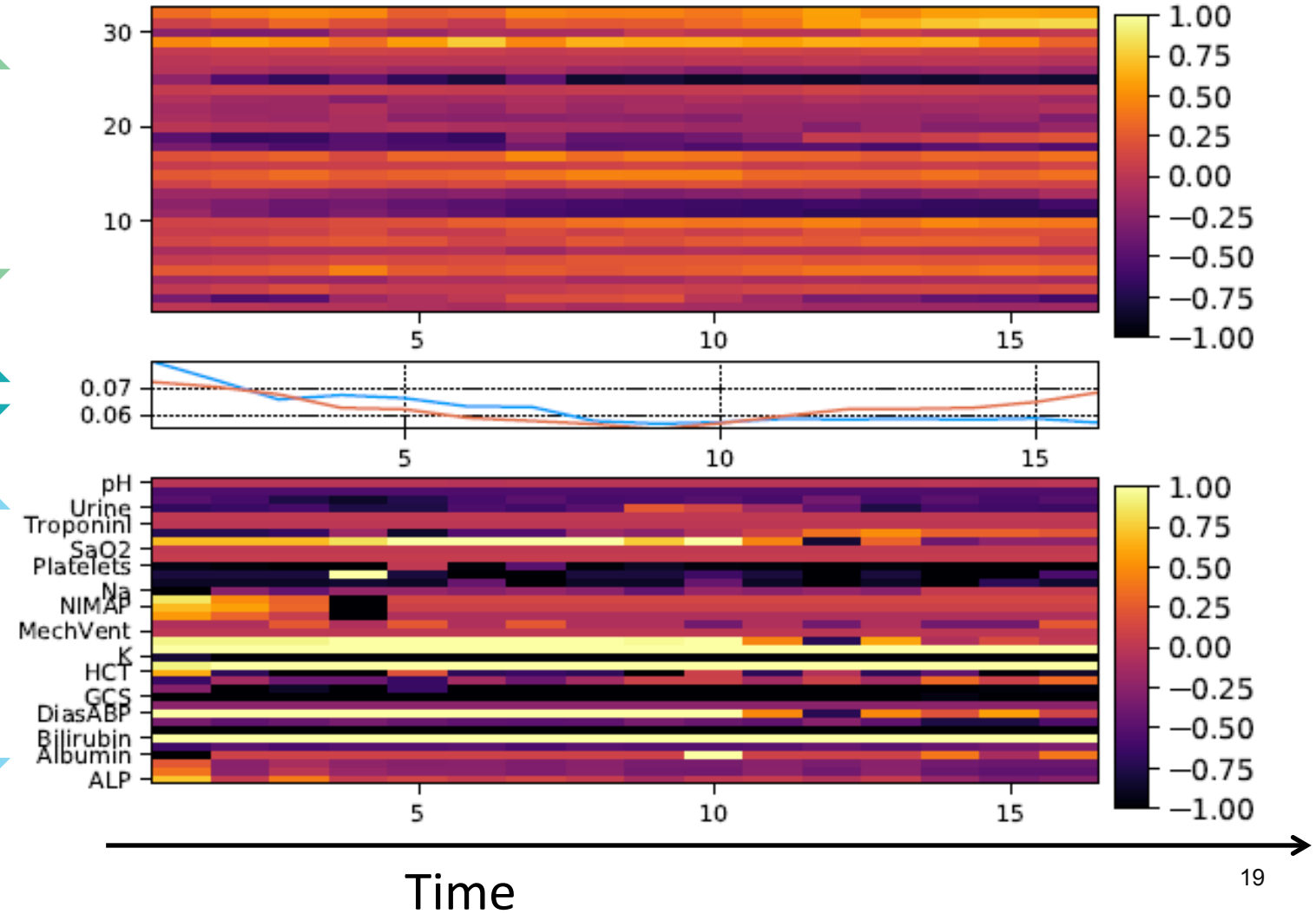
Attend to risks in ICU

#REF: Phuoc Nguyen, Truyen Tran, Svetha Venkatesh, "Deep Learning to Attend to Risk in ICU", *IJCAI'17 Workshop on Knowledge Discovery in Healthcare II: Towards Learning Healthcare Systems* (KDH 2017).

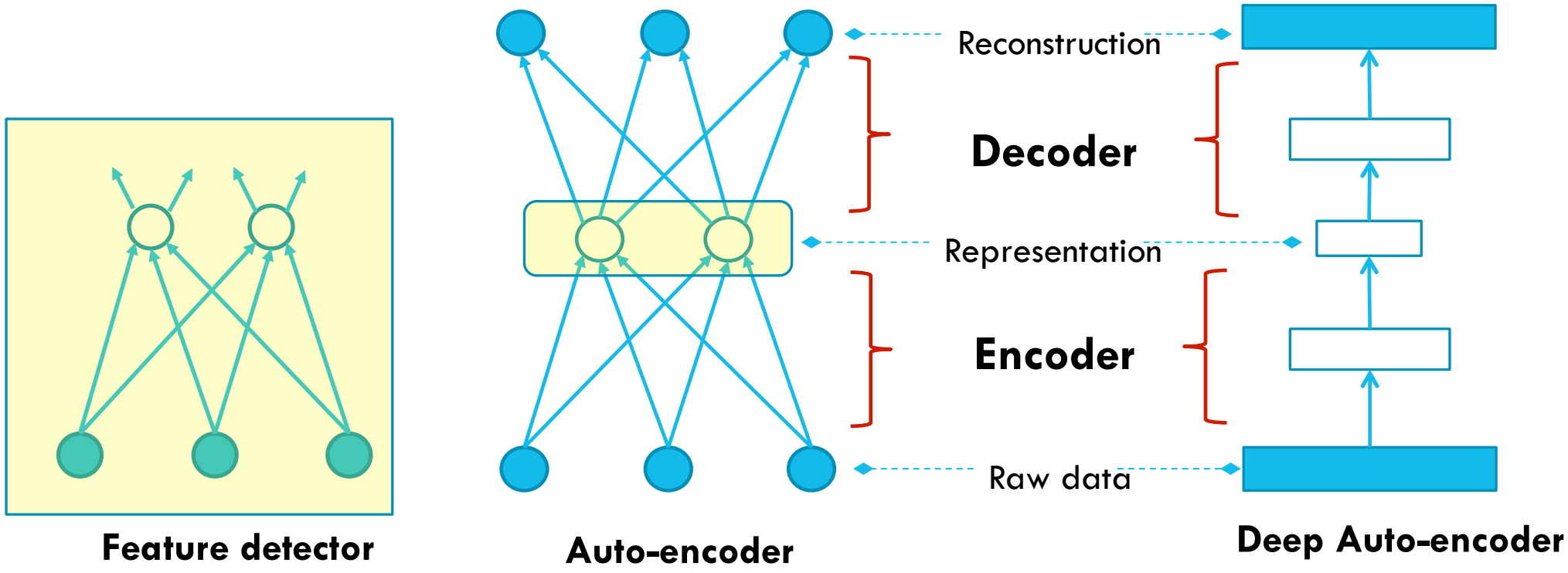
State transition

Attention probabilities

Physiological measures



DeepPatient: Representing medical records with Stacked Denoising Autoencoder



#Ref: Miotto, Riccardo, et al. "Deep patient: An unsupervised representation to predict the future of patients from the electronic health records." *Scientific reports* 6 (2016): 26094.

DeepPatient: Results on disease classification

Time Interval = 1 year (76,214 patients)			
		Classification Threshold = 0.6	
Patient Representation	AUC-ROC	Accuracy	F-Score
RawFeat	0.659	0.805	0.084
PCA	0.696	0.879	0.104
GMM	0.632	0.891	0.072
K-Means	0.672	0.887	0.093
ICA	0.695	0.882	0.101
DeepPatient	**0.773***	**0.929***	**0.181***

Trajectories modeling: Challenges & opportunities

Long-term dependencies

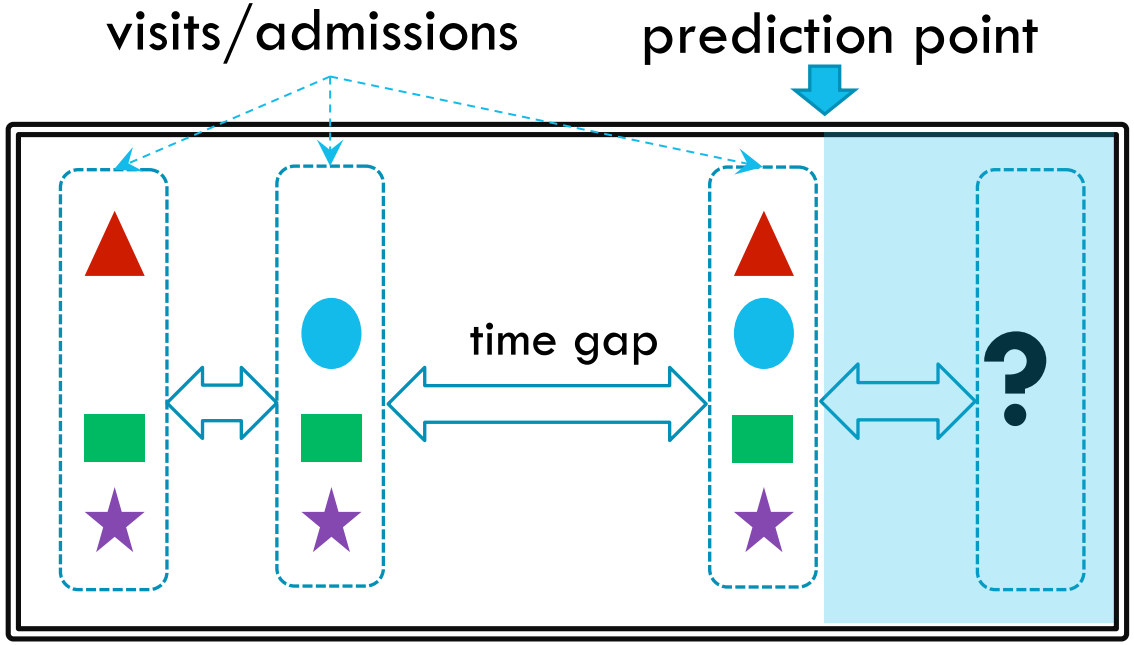
Irregular timing

Mixture of discrete codes and continuous measures

Complex interaction of diseases and care processes

Cohort of interest can be small (e.g., <1K)

Rich domain knowledge & ontologies



Multimodalities: Text, physiological signals (e.g., EEG/ECG), images (e.g., MRI, X-ray, retina), genomics

New modalities: social medial, wearable devices

Explainability!

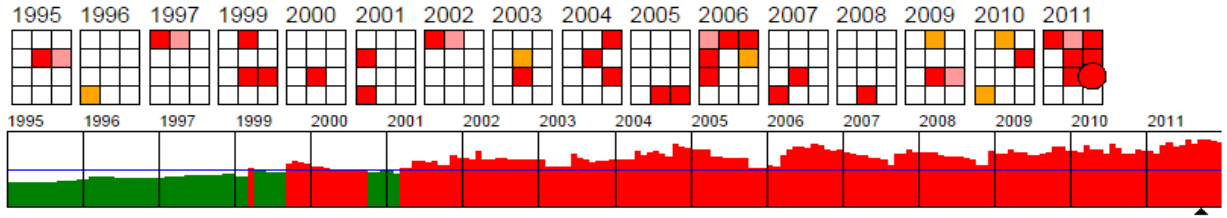
Untitled Document | prada1.it.deakin.edu.au:30

prada1.it.deakin.edu.au:3000/view/StrokeMap/000005

Project name | Home | Page1 | Page2

UR	000005
DOB	1936-01-01
Gender	Female
Occupation	home duties
Marital Status	Married
Risk	0.88 (2011/09/01)

History



Predictive Factors | All Factors

Disease	Other cataract
	Strep & staph cause dis class oth chptr
	Diverticular disease of intestine
	Oth symptoms signs inv cogn fn awareness
	Chronic kidney disease
	Unspecified urinary incontinence
	Essential (primary) hypertension
	Other disorders of urinary system
	Type 2 diabetes mellitus
	Heart failure
	Abnormalities of gait and mobility
	Pneumonia organism unspecified
	Oth sym signs inv nervous & M/S systems
	Malaise and fatigue
	Disrd lipoprotein metab & oth lipidaemia
	Atrial fibrillation and flutter
Admission	
pastProcNo	
Procedure	Generalised allied health interventions
	Conduction anaesthesia
	Cerebral anaesthesia
Emergency	
Context	Place of occurrence
	Personal history of medical treatment
Comorbidity	hypertension-uncomplicated
	diabetes-complicated
	cardiac-arrhythmias
pastRareProcNo	

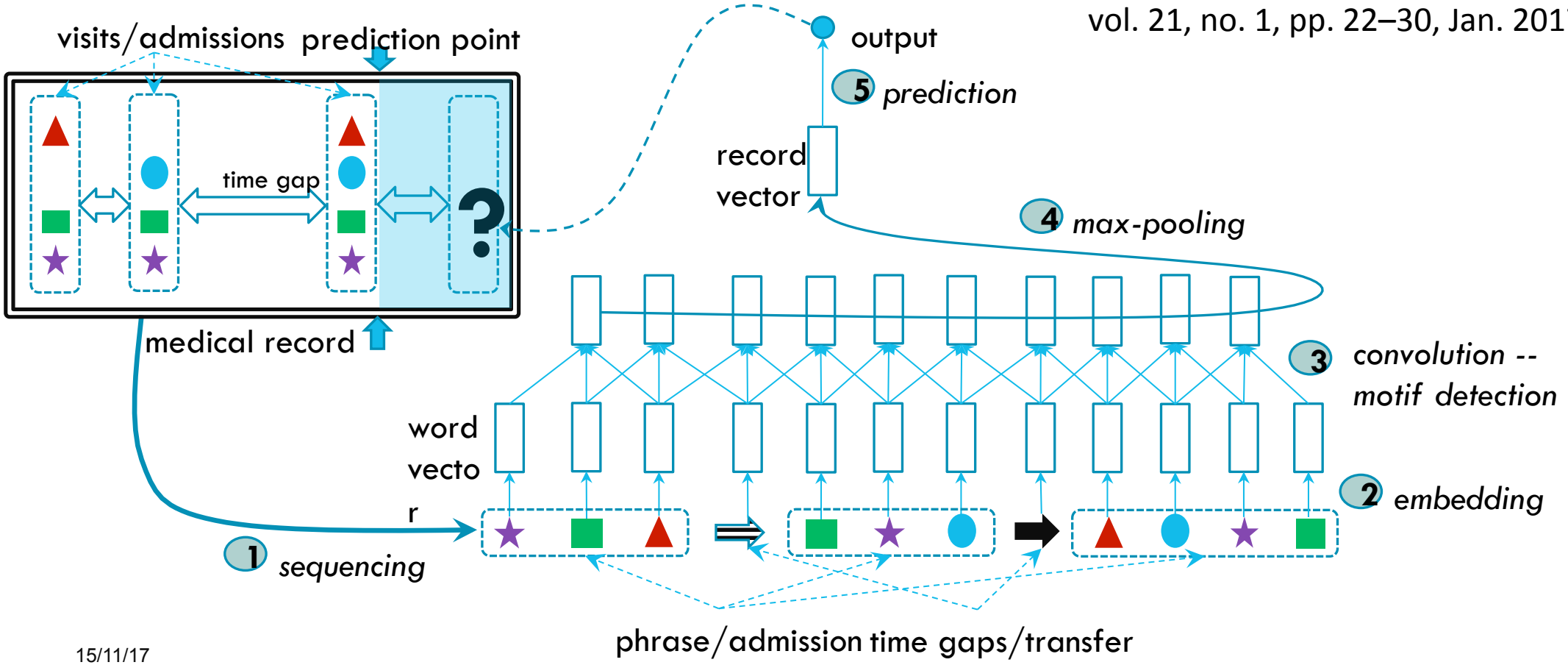
Events

1995/05/24	Emergency Admission (9.8 days)
59010	acute pyelonephritis
03842	septicemia due to other gramnegative or
5929	urinary calculus unspecified
4011	benign essential hypertension
4140	coronary atherosclerosis
8773	intravenous pyelogram

Visualisation and interpretation are keys!

Deepr: CNN for repeated motifs and short sequences

#REF: Phuoc Nguyen et al., Deepr: A Convolutional Net for Medical Records, *IEEE Journal of Biomedical and Health Informatics*, vol. 21, no. 1, pp. 22–30, Jan. 2017



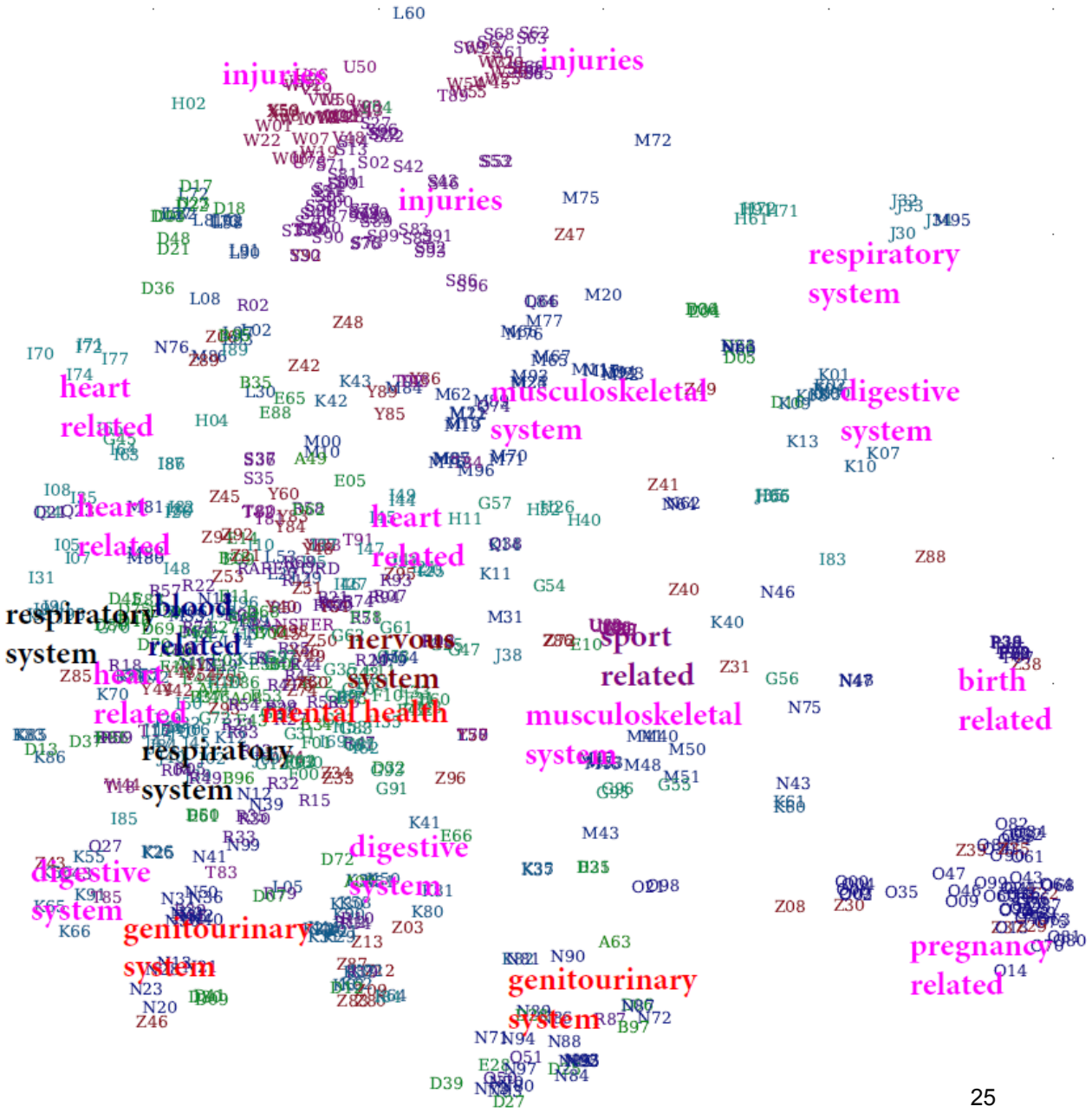
Deepr: Disease embedding & motifs detection

E11 I48 I50

Type 2 diabetes mellitus
Atrial fibrillation and flutter
Heart failure

E11 I50 N17

Type 2 diabetes mellitus
Heart failure
Acute kidney failure



DeepCare: intervened long-term memory of health

Illness states are a dynamic memory process → moderated by time and intervention

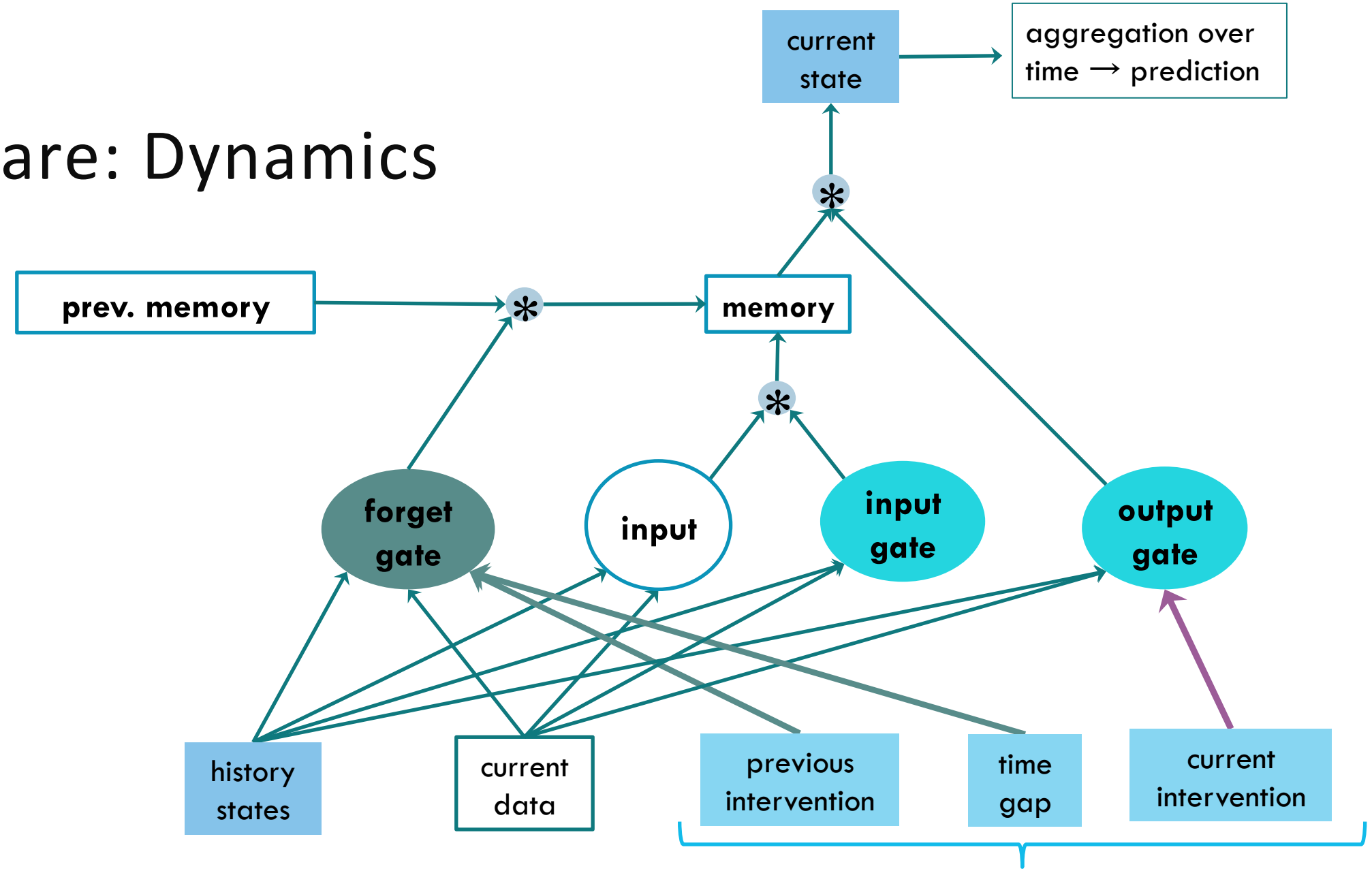
Discrete admission, diagnosis and procedure → vector embedding

Time and previous intervention → "forgetting" of illness

Current intervention → controlling the risk states

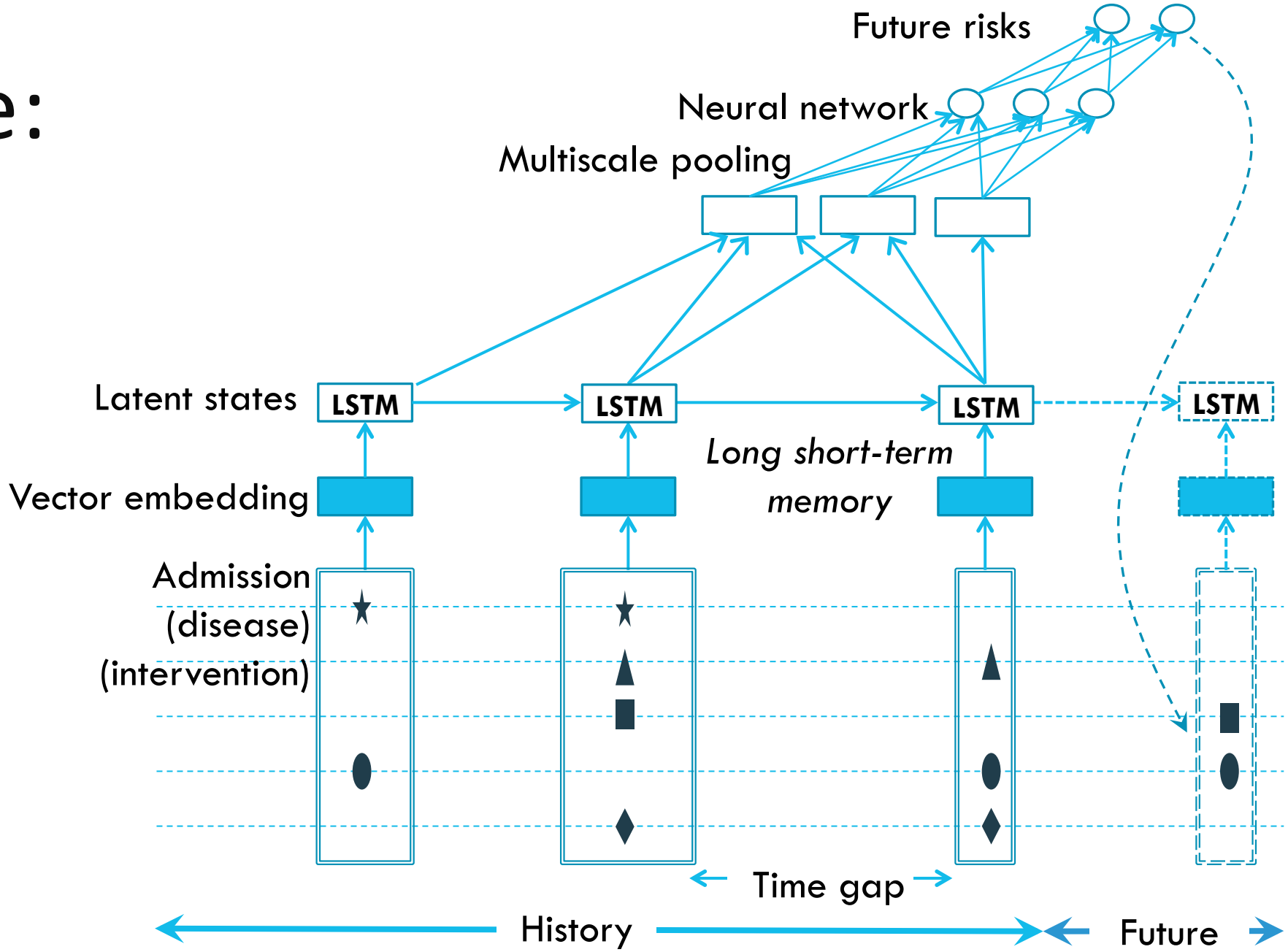
#REF: Trang Pham, et al., Predicting healthcare trajectories from medical records: A deep learning approach, *Journal of Biomedical Informatics*, April 2017, DOI: 10.1016/j.jbi.2017.04.001.

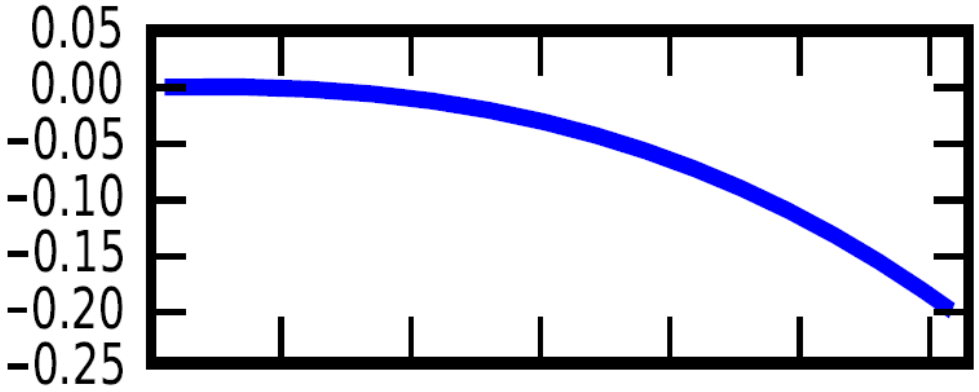
DeepCare: Dynamics



New in DeepCare

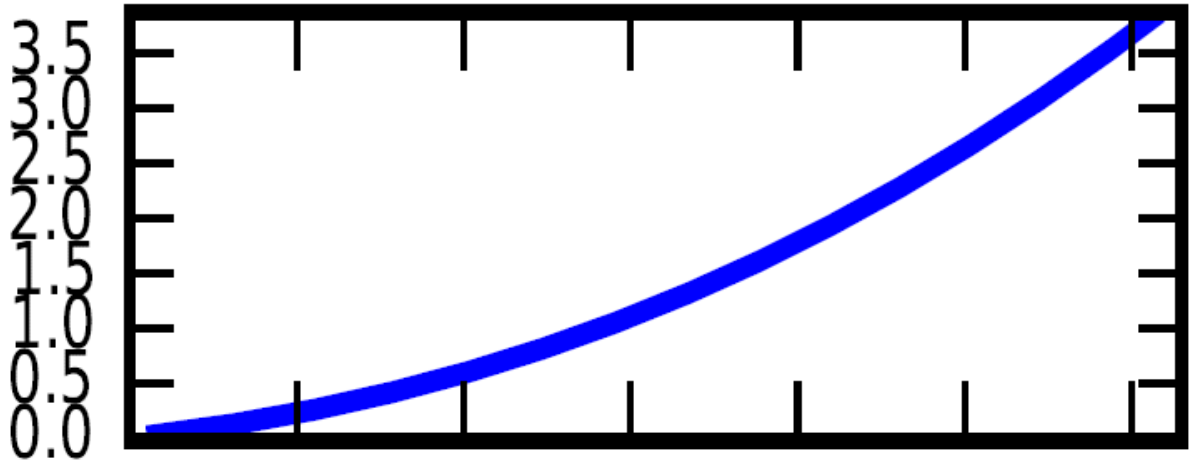
DeepCare: Structure





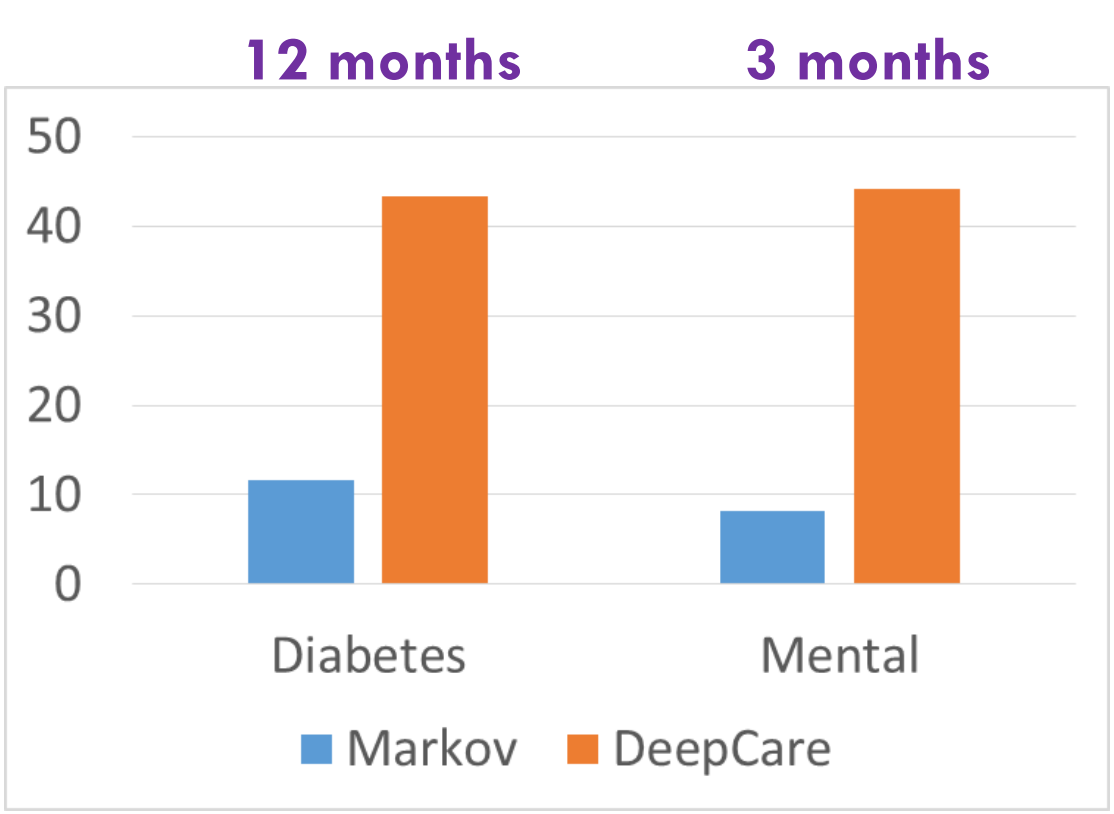
→ decreasing illness

→ Increasing illness

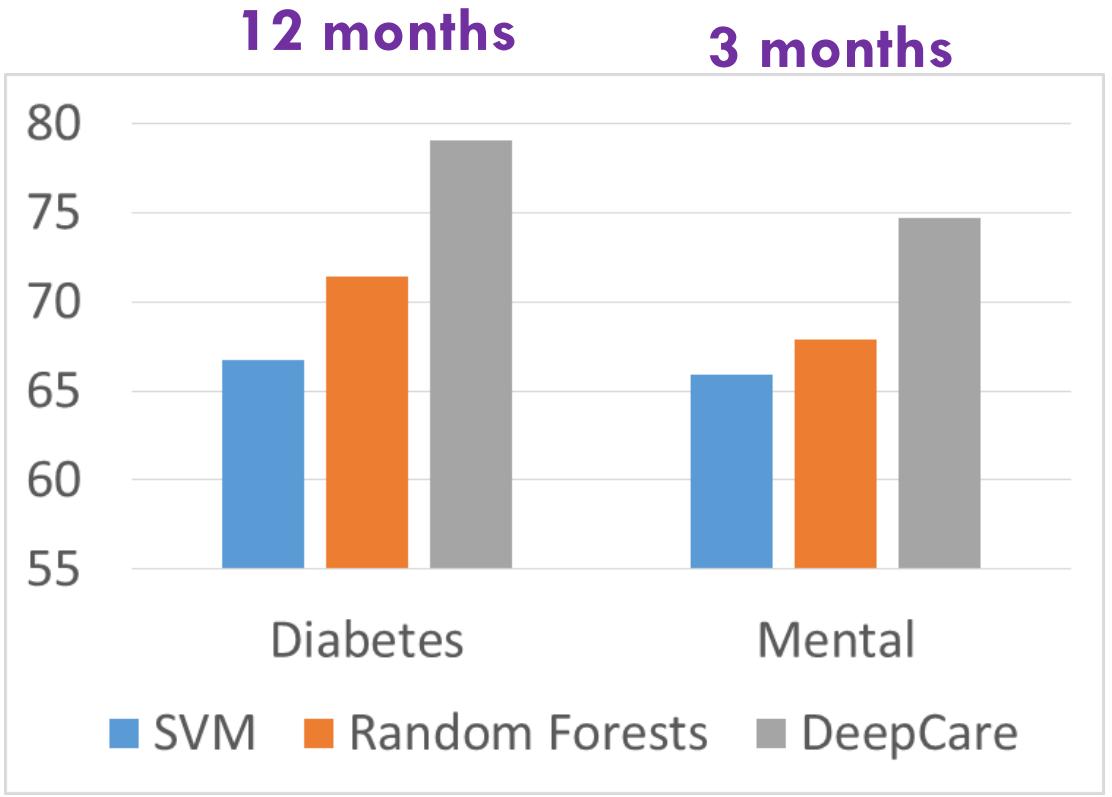


DeepCare: Two modes of forgetting as a function of time

DeepCare: prediction results



Intervention recommendation (precision@3)



Unplanned readmission prediction (F-score)

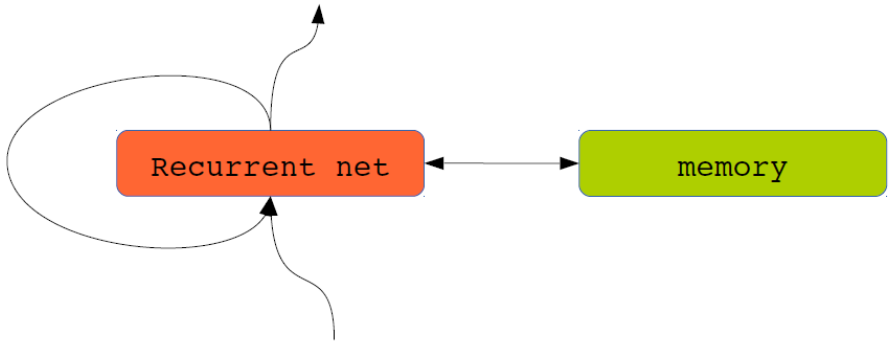
Trajectories prediction

Generating a subset of treatments

Generating an entire health/care trajectory

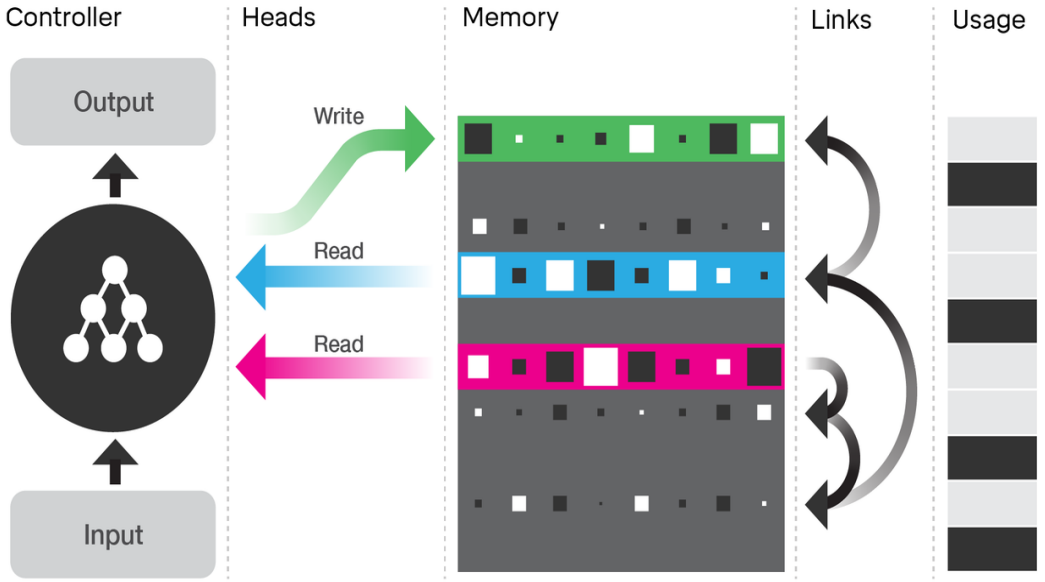
Challenges: global loss, meaningful evaluation metrics

A solution: Memory-augmented neural nets (MANN)



(LeCun, 2015)

Illustration of the DNC architecture



Source: deepmind.com

#REF: Graves, Alex, et al. "Hybrid computing using a neural network with dynamic external memory." *Nature* 538.7626 (2016): 471-476.

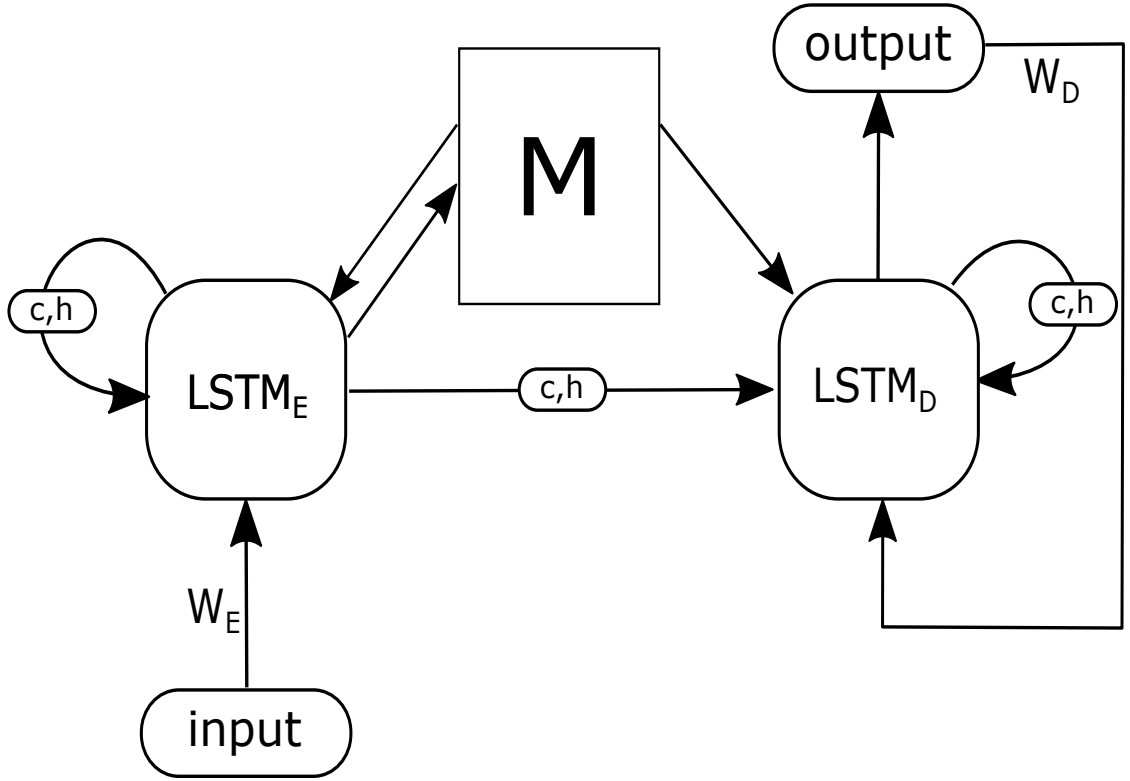
Dual Controller Write-protected Memory-Augmented Neural Networks (DCw-MANN) *(Le et al, work in progress)*

Separate controllers for encoding and decoding

Memory is write-protected during decoding

Use memory as a replacement for:

- Attention
- Skip connection



DCw-MANN

Model	Procedure Output		Drug Output	
	Recall	BLEU	Recall	BLEU
Logistic Regression	25.60	N/A	41.21	N/A
Random Forest	26.84	N/A	49.06	N/A
Seq2Seq	27.16	36.42	23.56	37.71
Seq2Seq with attention	27.33	36.83	24.82	39.70
DNC	28.28	37.69	58.16	61.72
DC-MANN	28.54	38.51	58.79	62.70
DCw-MANN	**30.36**	**39.74**	**59.76**	**63.80**

Table 3. Results on MIMIC-III dataset for procedure prediction and drug prescription. The higher the better.

Modeling multiple disease-treatment interactions over time

Co-morbidity is the norm in modern medicine

Each hospital visit contains a set of diseases and a set of treatments

There are interactions between multi-diseases and multiple-treatments

Algebraic view: **Health = *RNN*(Illness – Intervention)**

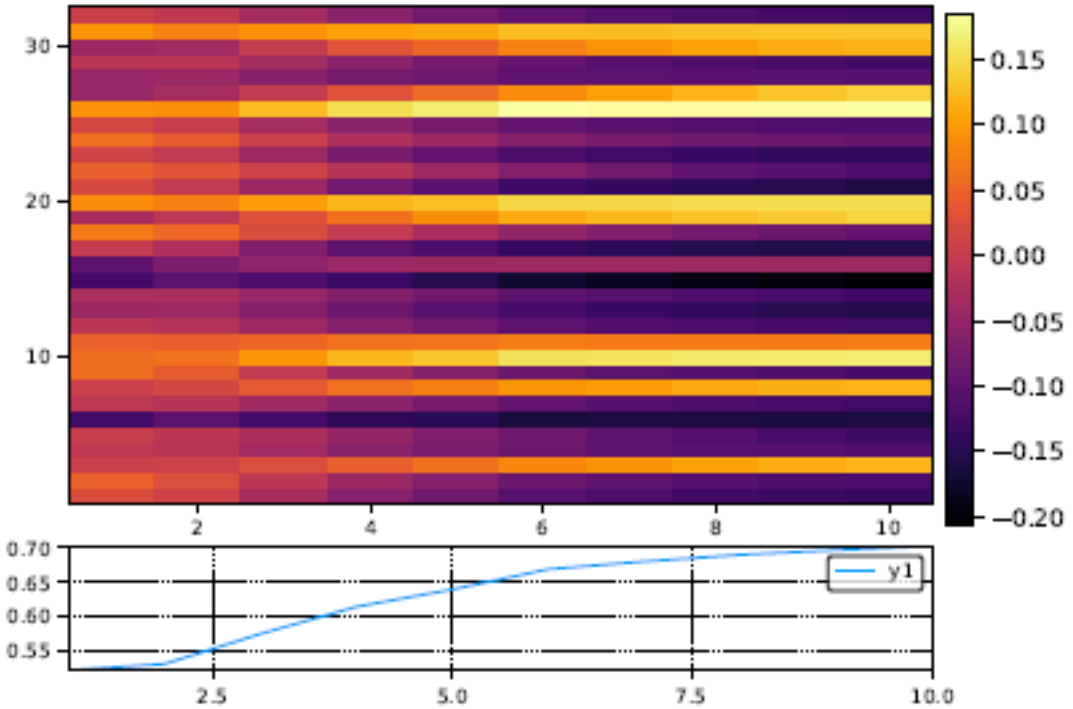
$$\boldsymbol{v}_t = \rho\left(\Delta\right) \quad \text{where} \quad \Delta = \boldsymbol{d}_t - \boldsymbol{p}_t$$

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

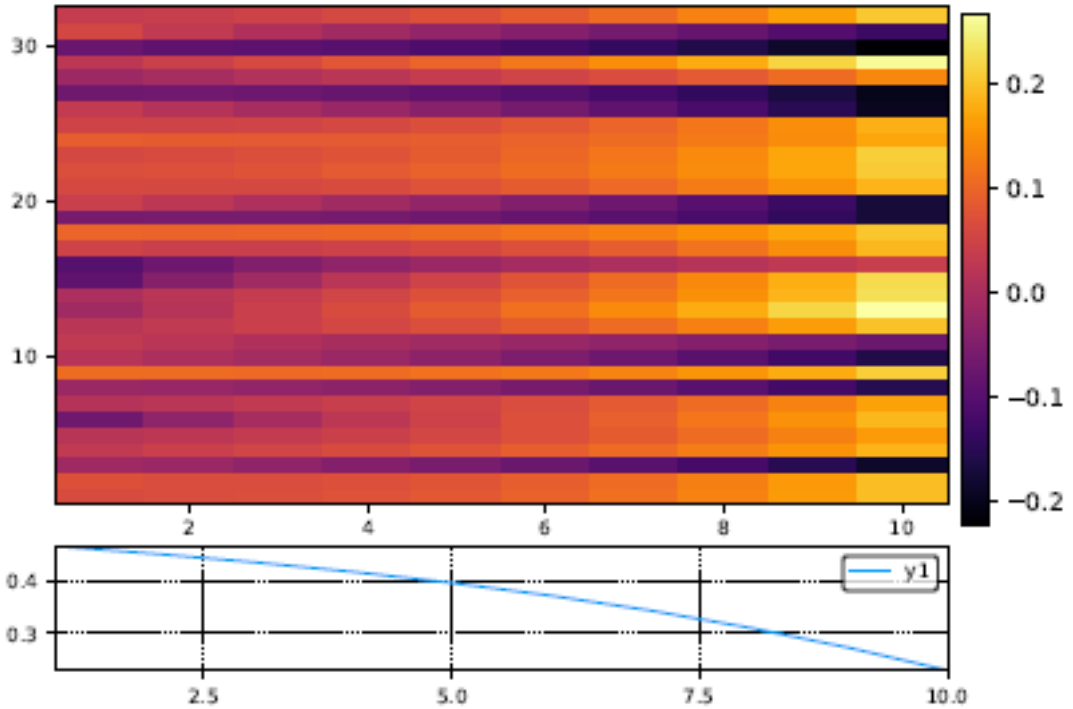
#REF: P Nguyen, T Tran, S Venkatesh, "Finding Algebraic Structure of Care in Time: A Deep Learning Approach", *NIPS Workshop on Machine Learning for Health (ML4H)*, 2017

Results (AUC)

Method	*Diabetes*	*Mental health*
BoW+LR	0.673	0.705
Deepr [14]	0.680	0.714
MDMTP+LTSM	**0.718**	**0.726**
MDMT+LSTM	**0.701**	**0.730**



(a) Worsening progression ($P = 0.70$)



(b) Improving progression ($P = 0.23$)

Big room: Towards personalized healthcare

Medical practice as recommender systems (Xavier Amatriain, healthcare Recsys Workshop, Como, 2017)

Clinical Practice Guides are not personalized

Research done on “homogeneous”, healthy subjects

It is very hard for doctors to “manually” personalize their “recommendations”

Other refs

Choi, Edward, et al. "Doctor ai: Predicting clinical events via recurrent neural networks." *Machine Learning for Healthcare Conference*. 2016.

Lipton, Zachary C., et al. "Learning to diagnose with LSTM recurrent neural networks." *arXiv preprint arXiv:1511.03677*(2015).

Harutyunyan, Hrayr, et al. "Multitask Learning and Benchmarking with Clinical Time Series Data." *arXiv preprint arXiv:1703.07771* (2017).

Choi, Edward, et al. "RETAIN: An interpretable predictive model for healthcare using reverse time attention mechanism." *Advances in Neural Information Processing Systems*. 2016.

Agenda

Topic 1: Introduction (20 mins)

Topic 2: Brief review of deep learning (25 mins)

- Classic architectures
- Capsules
- Graphs
- Memory-augmented nets

Topic 3: Genomics (25 mins)

- Nanopore sequencing
- Genomics modelling

QA (10 mins)

Break (20 mins)

Topic 4: Biomedical imaging (15 mins)

- Cellular imaging
- Diagnostics imaging
- EEG/ECG

Topic 5: Healthcare (25 mins)

- Time-series of physio measures
- Trajectories prediction

Topic 6: Generative biomed (30 mins)

- Few-shot learning
- Generative models
- Drug design
- Future outlook

QA (10 mins)

Few-shot deep learning

Lots of biomedical problems are data poor

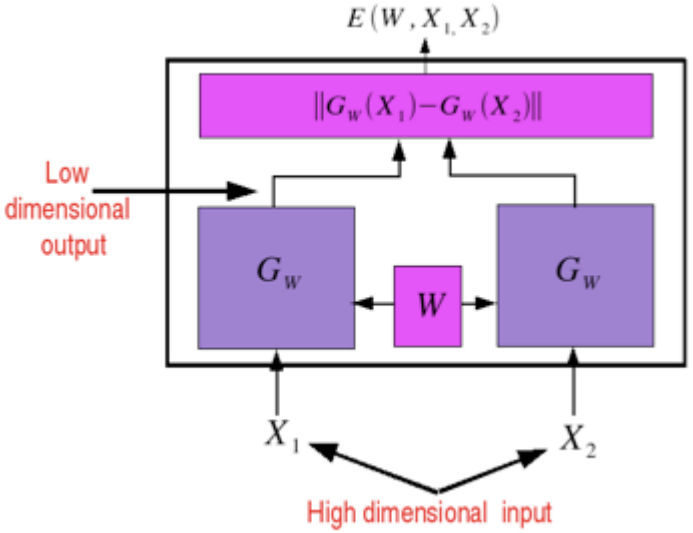
- Rare drugs
- Rare diseases

Distance metrics learning (DML) methods

- Learn to full any pair of the similar data points, and push the dissimilar
- Well-known methods: Siamese networks

Meta-learning strategies

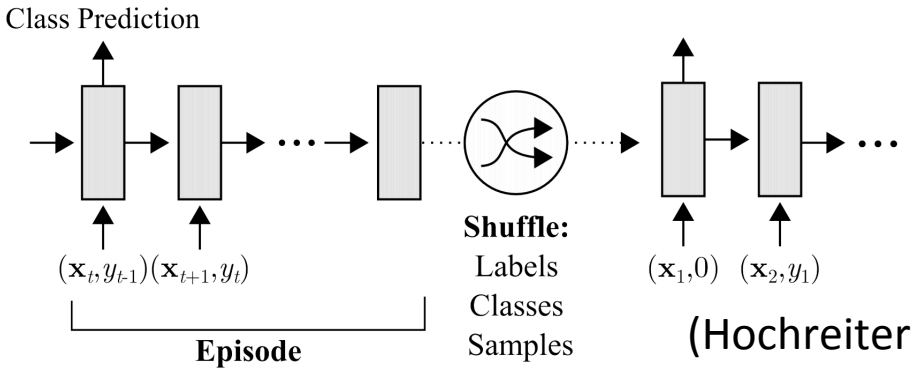
- Tasks are presented in sequence
- New tasks can borrow from similar prior tasks



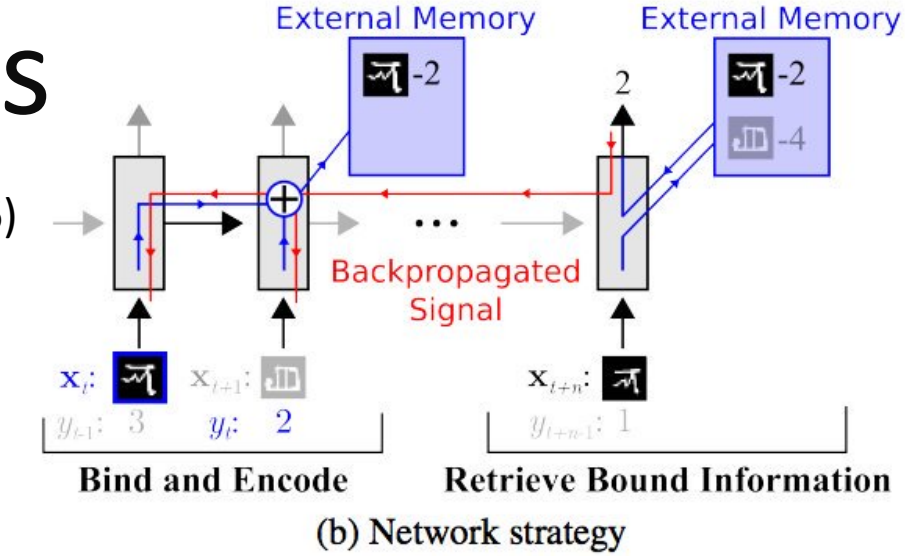
#REF: Chopra, Sumit, Raia Hadsell, and Yann LeCun. "Learning a similarity metric discriminatively, with application to face verification." *Computer Vision and Pattern Recognition, 2005. CVPR 2005. IEEE Computer Society Conference on*. Vol. 1. IEEE, 2005.

Meta-learning strategies

(Santoro et al, 2016)



(Hochreiter et al., 2001)



(Mishra et al, 2017)

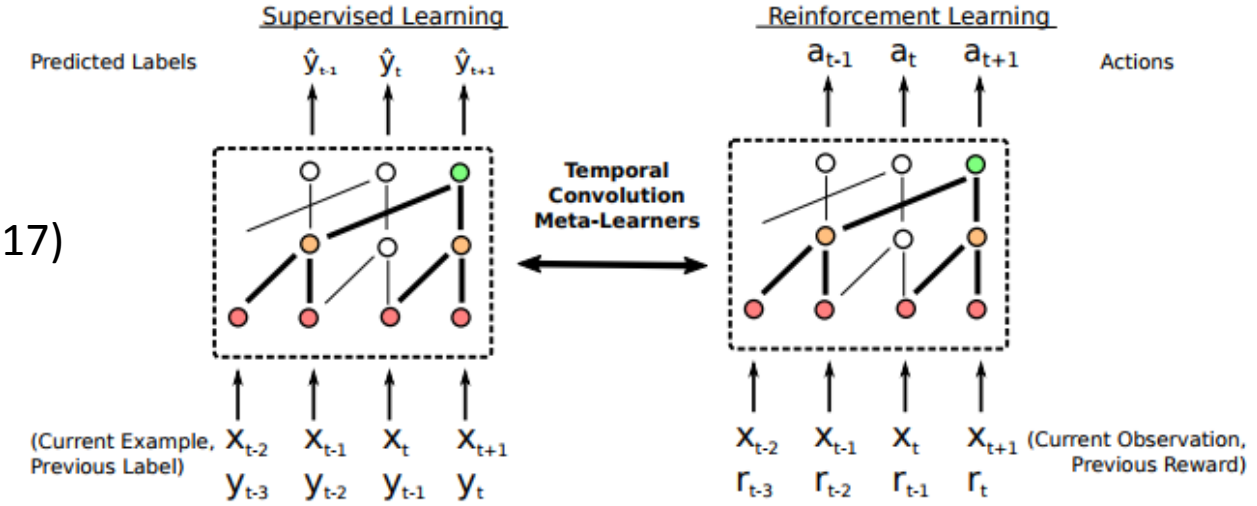


Figure 1: Overview of our temporal-convolution-based meta-learner (TCML). The same class of model architectures can be applied to both supervised and reinforcement learning.

Generative models

Many applications:

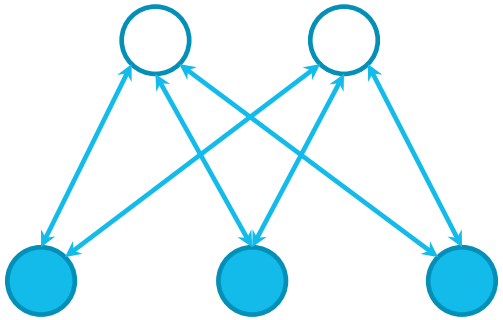
- Text to speech
- **Simulate data that are hard to obtain/ share in real life (e.g., healthcare)**
- Generate meaningful sentences conditioned on some input (foreign language, image, video)
- Semi-supervised learning
- Planning

$$\mathbf{v} \sim P_{model}(\mathbf{v})$$
$$P_{model}(\mathbf{v}) \approx P_{data}(\mathbf{v})$$

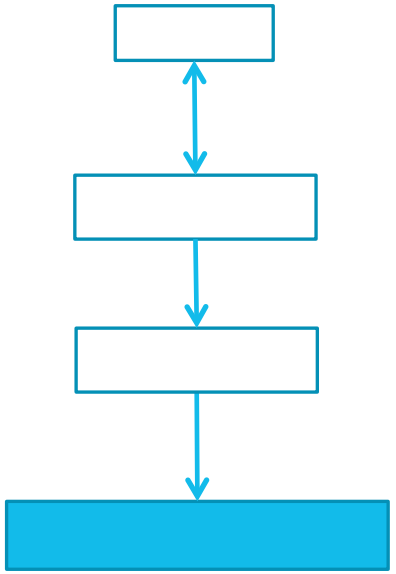
A family: RBM → DBN → DBM

$$p(\mathbf{v}, \mathbf{h}; \psi) \propto \exp\left[-E(\mathbf{v}, \mathbf{h}; \psi)\right]$$

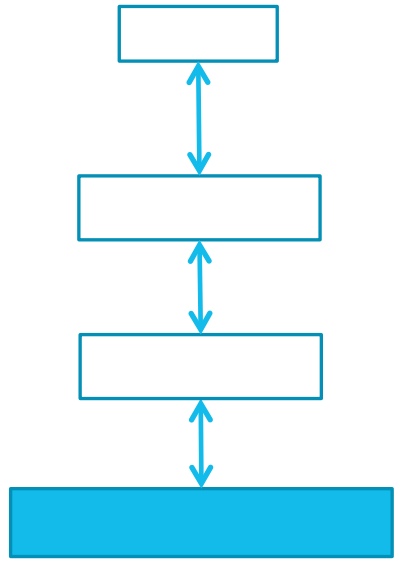
energy



Restricted Boltzmann Machine (~1994, 2001)



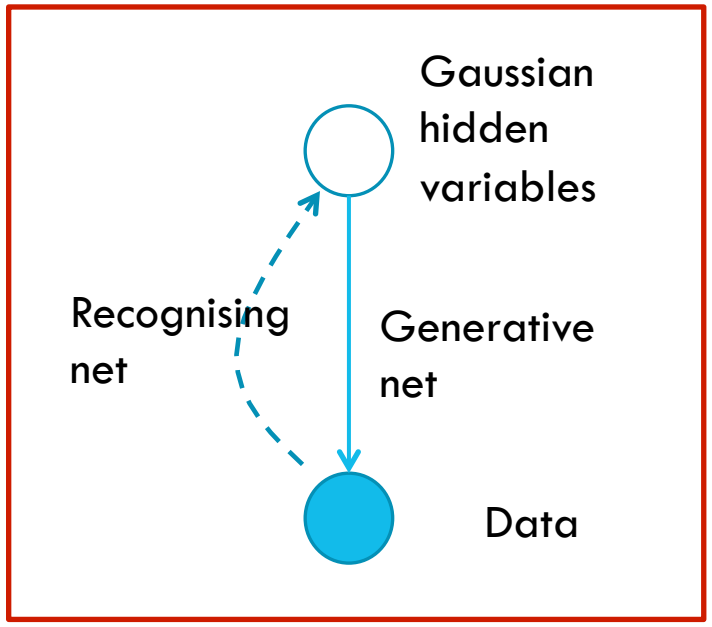
Deep Belief Net (2006)

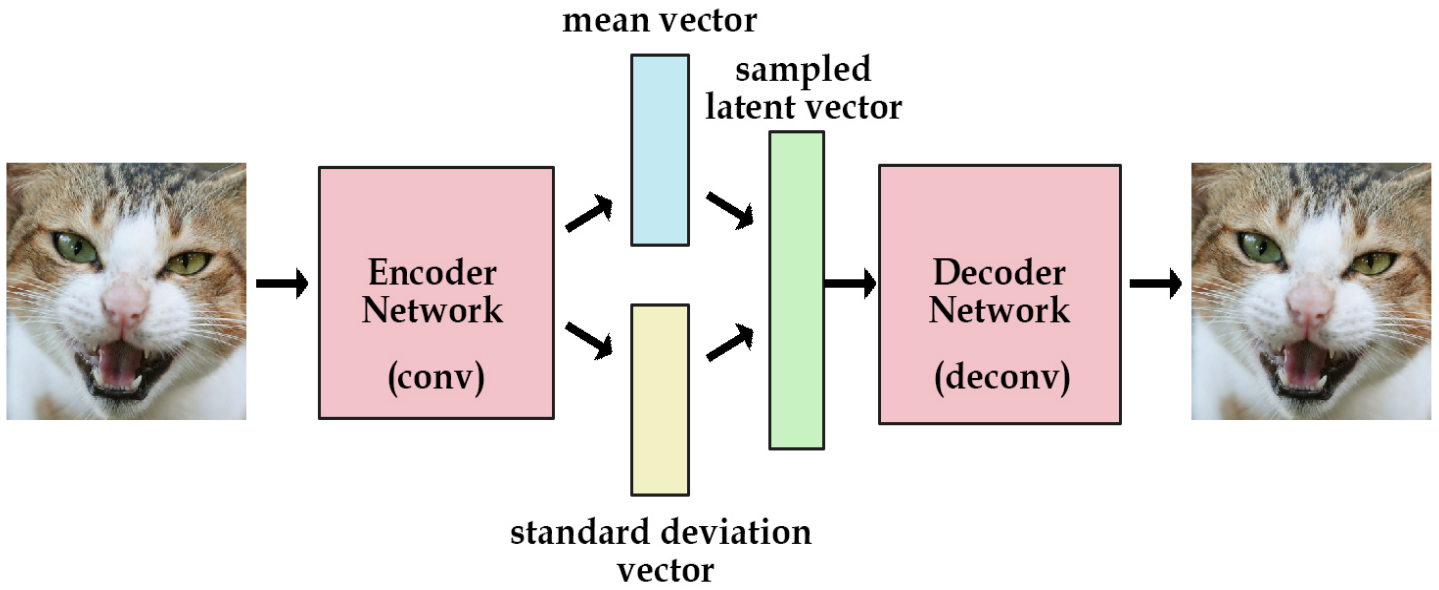


Deep Boltzmann Machine (2009)

Variational Autoencoder
(Kingma & Welling, 2014)

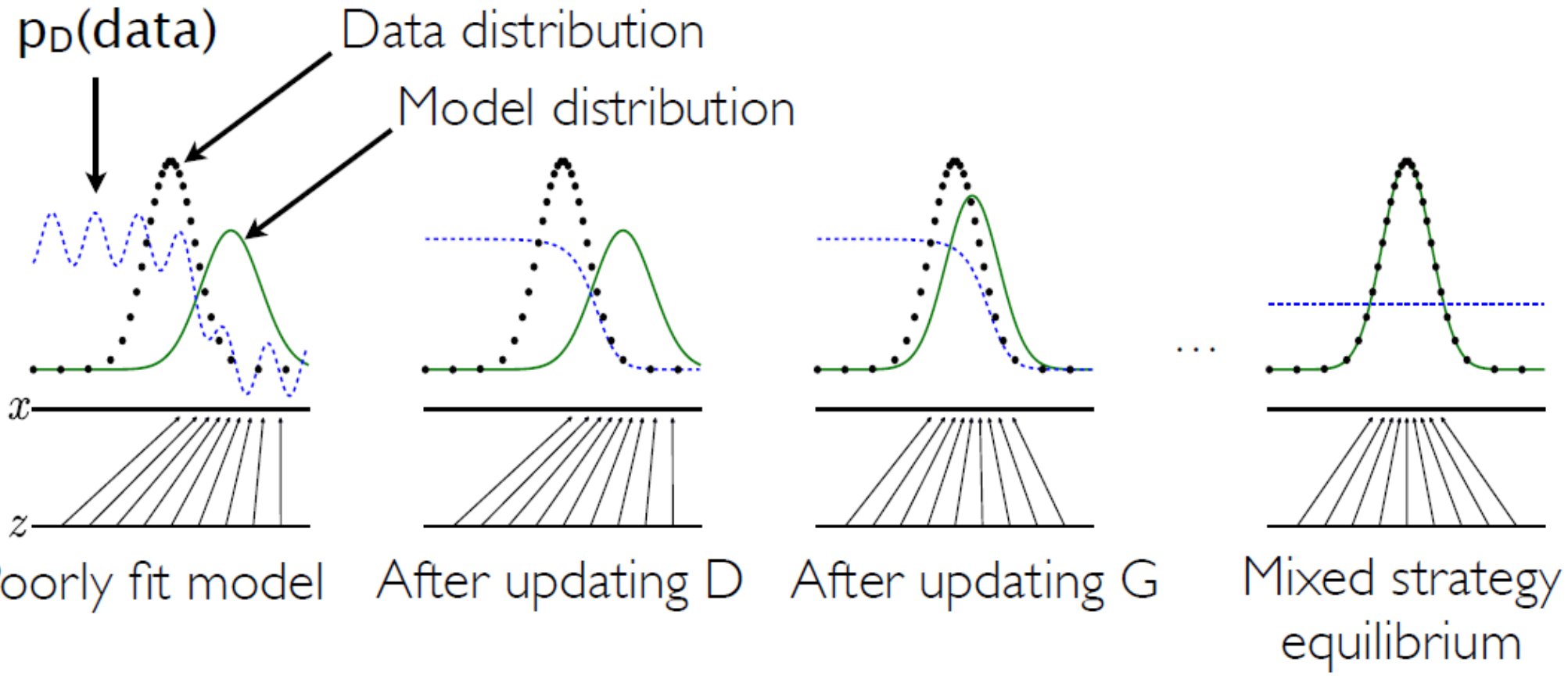
Two separate processes: generative (hidden → visible) versus recognition (visible → hidden)





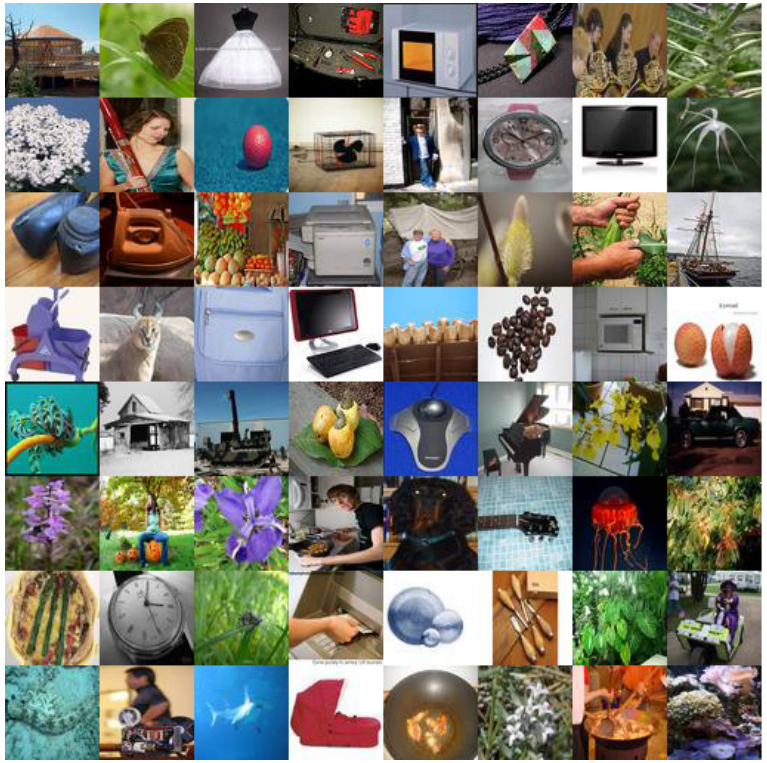
http://kvfrans.com/variational-autoencoders-explained/

GAN: implicit density models
(Adapted from Goodfellow's, NIPS 2014)

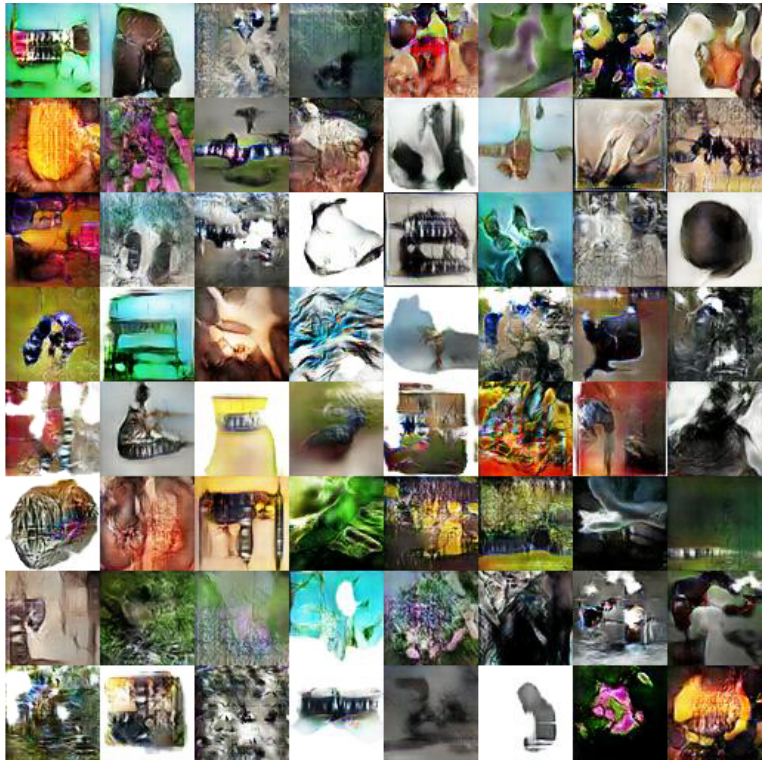


GAN: Generated Samples

Some high quality pictures generated by GAN



Real



Generated

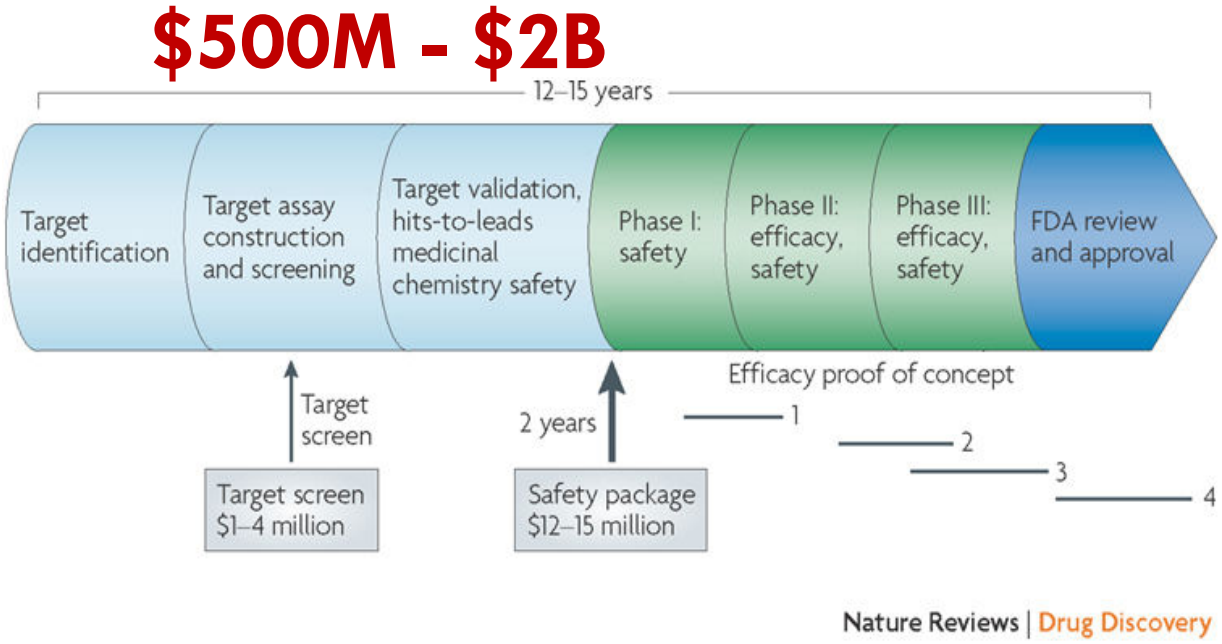
http://kvfrans.com/generative-adversial-networks-explained/

Generative deep learning for drug discovery

Predicting bioactivities from molecules

Drug representation, unsupervised learning from graphs

Generate from bioactivities to molecular graphs



#REF: Roses, Allen D. "Pharmacogenetics in drug discovery and development: a translational perspective." *Nature reviews Drug discovery* 7.10 (2008): 807-817.

Combinatorial chemistry

Generate variations on a template

Returns a list of molecules from this template that

- Bind to the pocket with good pharmacodynamics?
- Have good pharmacokinetics?
- Are synthetically accessible?

#REF: Talk by Chloé-Agathe Azencott titled “Machine learning for therapeutic research”, 12/10/2017

First step: Map molecule → drug properties (binding/acting)

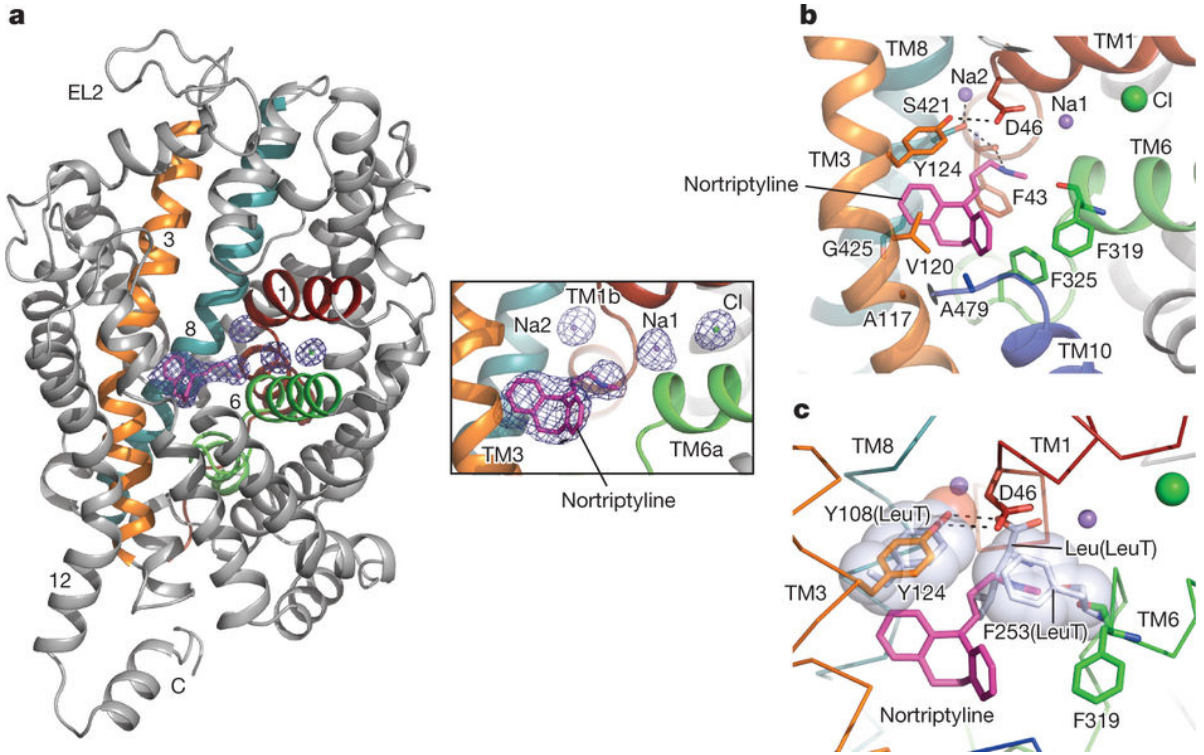
Drugs are small bio-molecules

Traditional techniques:

- Graph kernels (ML)
- Molecular fingerprints (Chemistry)

Modern techniques

- Molecule as graph: atoms as nodes, chemical bonds as edges

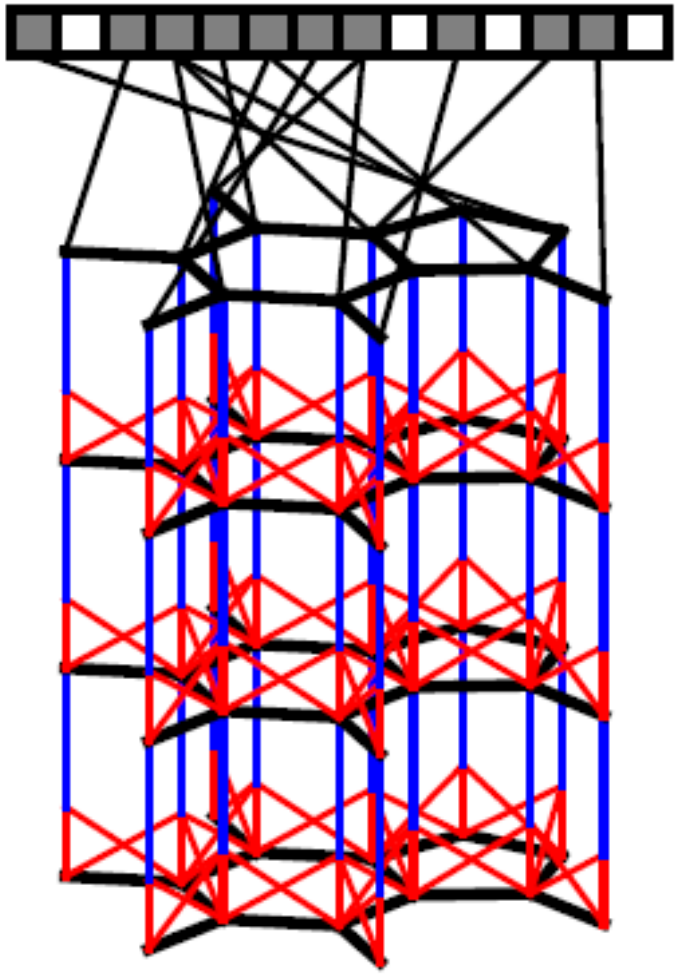


#REF: Penmatsa, Aravind, Kevin H. Wang, and Eric Gouaux. "X-ray structure of dopamine transporter elucidates antidepressant mechanism." *Nature* 503.7474 (2013): 85-90.

Molecular fingerprints

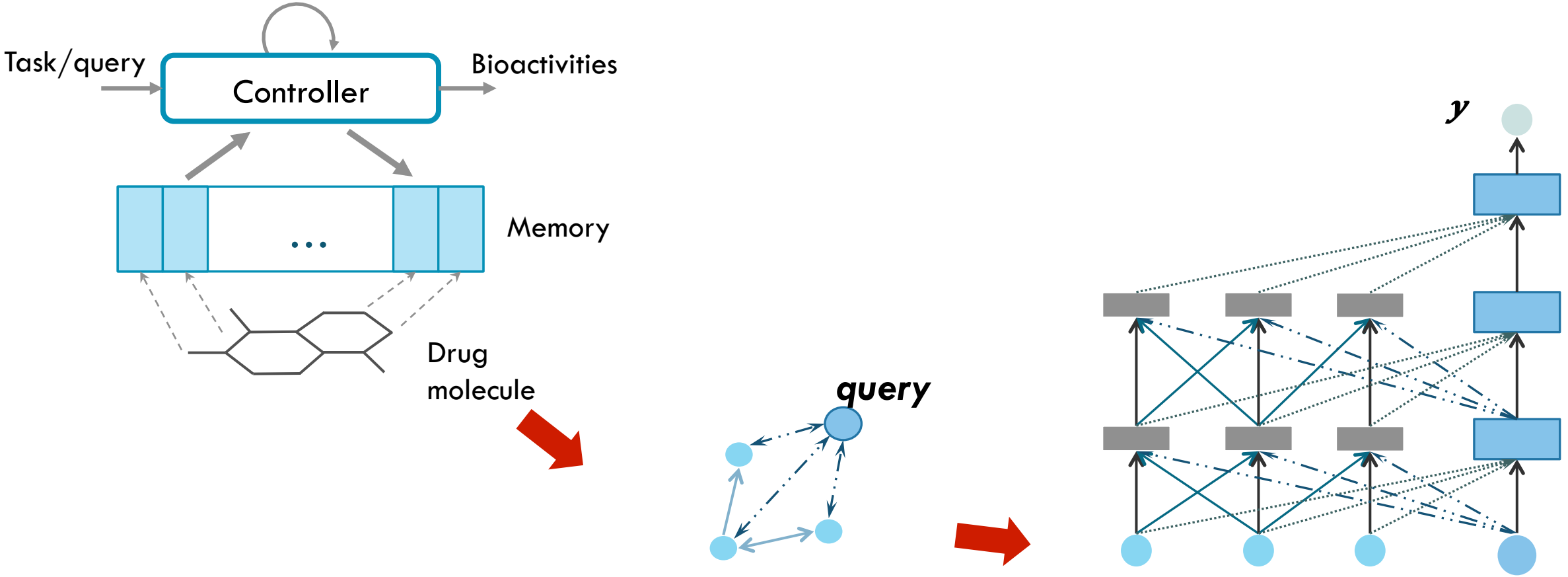
Algorithm 1 Circular fingerprints.

Input: molecule, radius R, fingerprint length S
Initialize: fingerprint vector $\boldsymbol{x} \leftarrow \mathbf{0}_S$
foreach atom a **in** molecule
    $\boldsymbol{r}_a \leftarrow q(a)$ #*extract initial atom features*
for $L = 1$ **to** R #*loop through layers*
    foreach atom a **in** molecule
        $\boldsymbol{r}_1 ... \boldsymbol{r}_N = \text{neighbors}(a)$
        $\boldsymbol{v} \leftarrow [\boldsymbol{r}_a, \boldsymbol{r}_1, ..., \boldsymbol{r}_N]$ #*combine neighbor features*
        $\boldsymbol{r}_a \leftarrow \text{hash}(\boldsymbol{v})$ #*refine atom features*
        $i \leftarrow \text{mod}(\boldsymbol{r}_a, S)$ #*convert to index*
        $x_i \leftarrow 1$ #*Write 1 (indicator) at index*
Return: binary vector $\boldsymbol{x}$.



#REF: Duvenaud, David K., et al. "Convolutional networks on graphs for learning molecular fingerprints." *Advances in neural information processing systems*. 2015.

Graph memory networks
(Pham et al, 2017, work in progress)



Graph memory networks: Results

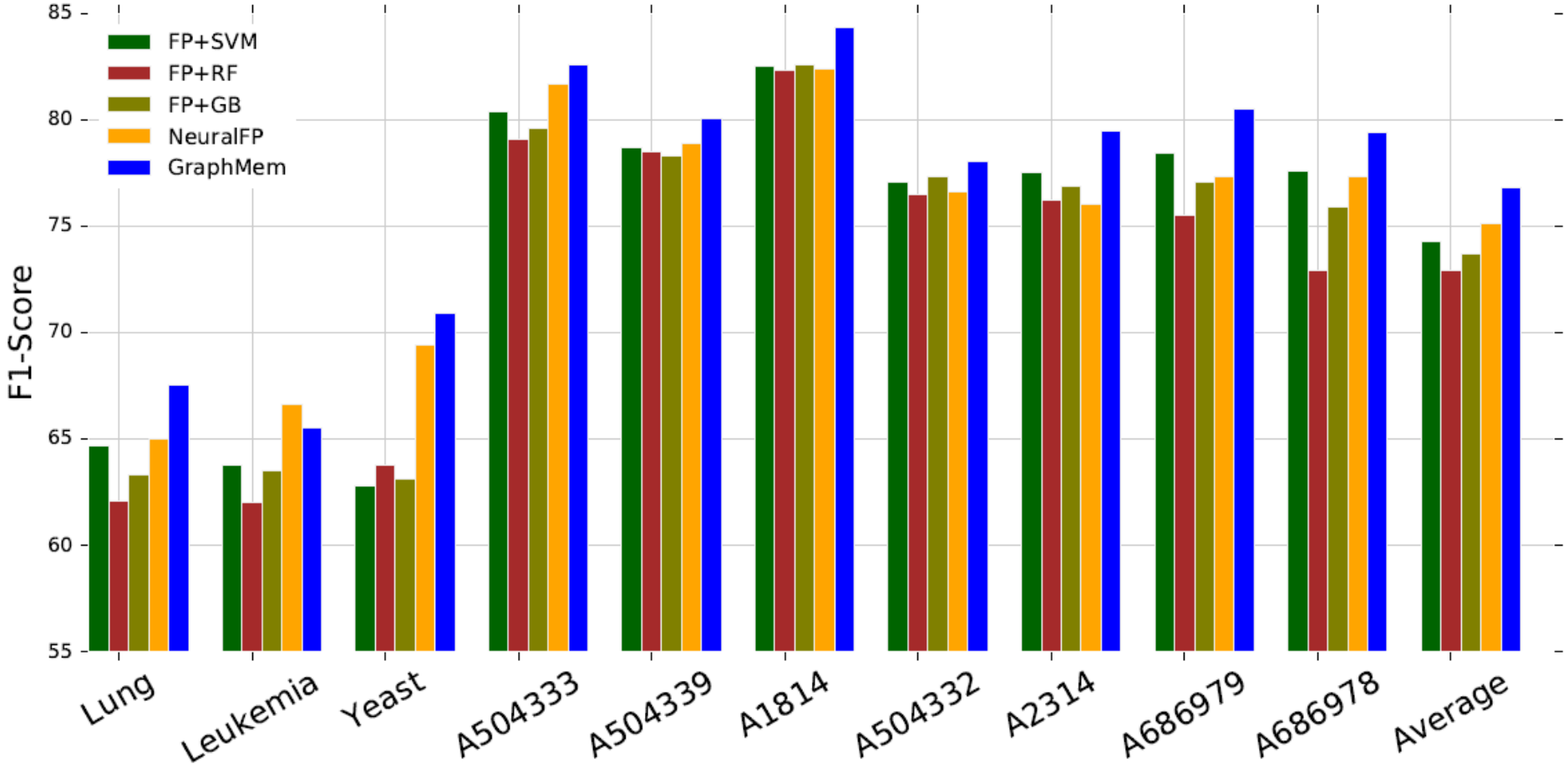
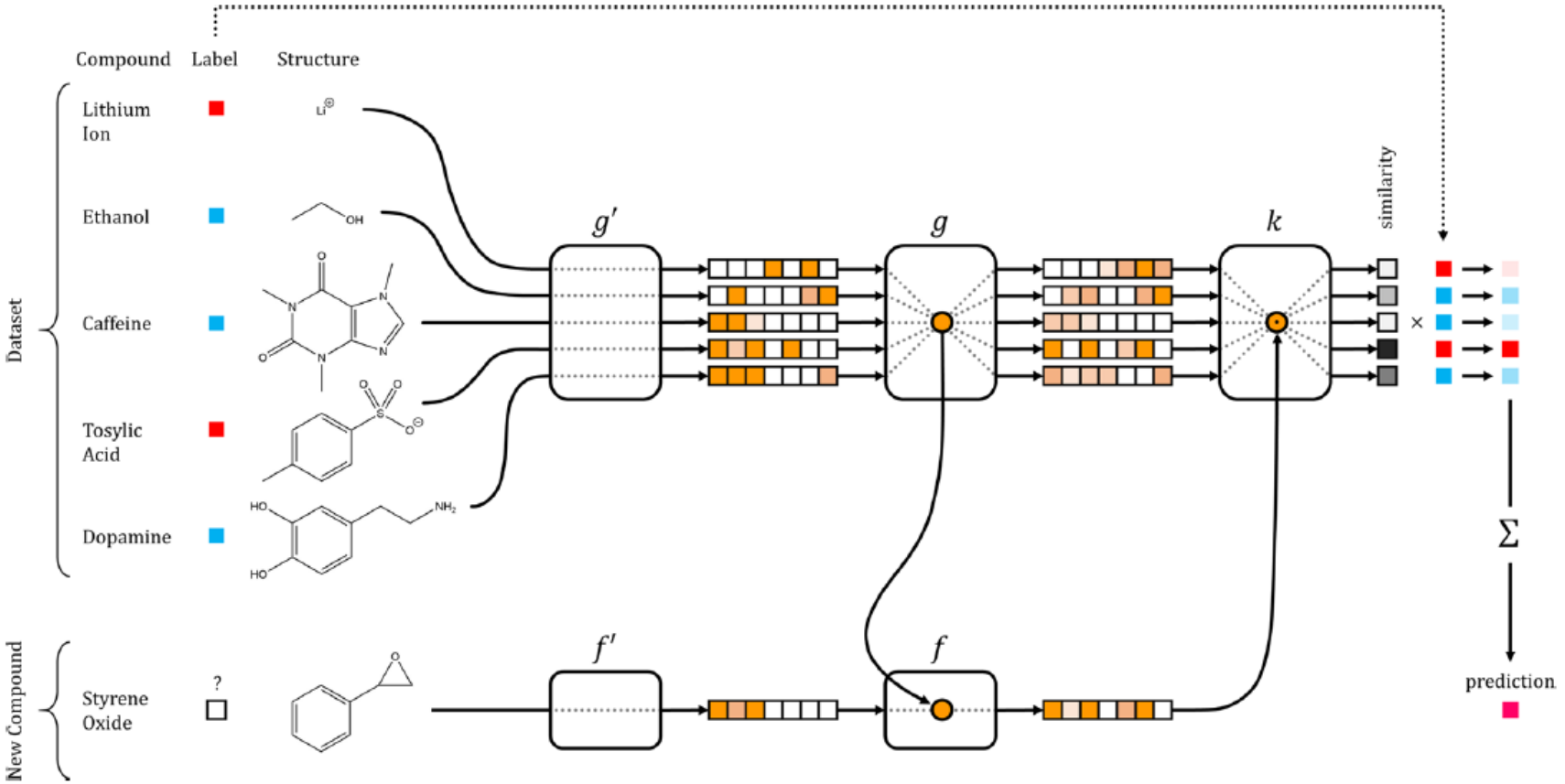


Figure 2: F1-score (%) for NCI datasets. FP = Fingerprint; RF = Random Forests; GBM = Gradient Boosting Machine. Best view in color.

One-shot learning for drug discovery



#REF: Altae-Tran, Han, et al. "Low Data Drug Discovery with One-Shot Learning." *ACS central science* 3.4 (2017): 283-293.

Drug design and generation

We now have methods for compute bioactivties of a drug molecule

We need a reverse method to generate drug molecules from desirable bioactivities

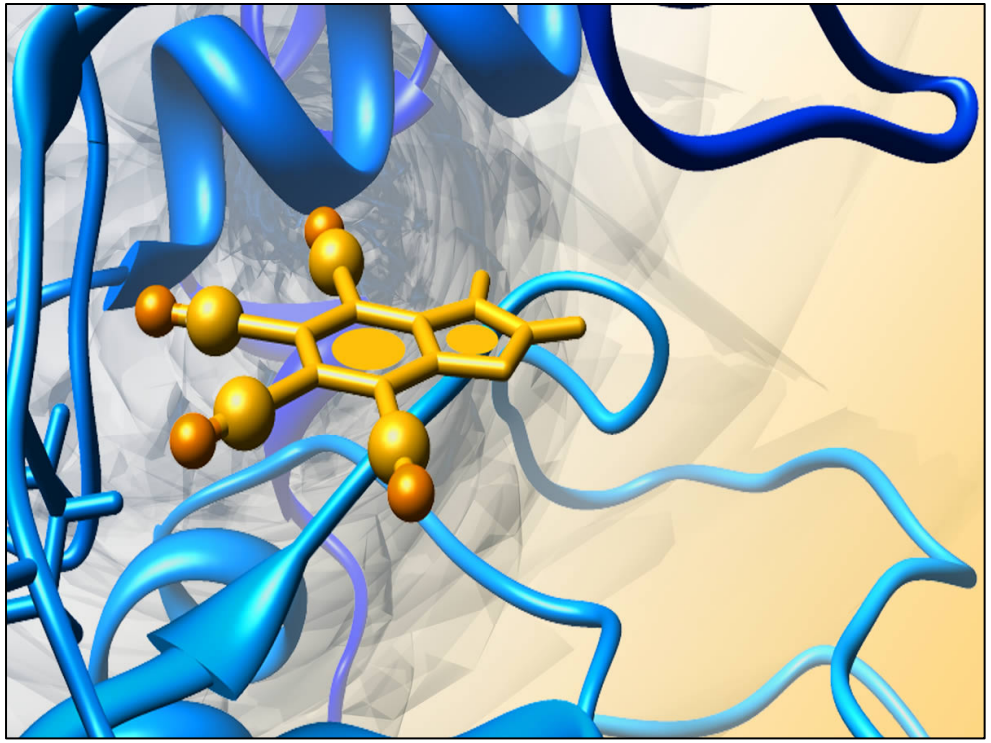
The space of drugs is estimated to be 1e+23 to 1e+60

- Only 1e+8 substances synthesized thus far.

It is impossible to model this space fully.

The current technologies are not ready for graph generations.

But approximate techniques do exist.



Source: pharmafactz.com

The trick: Molecule → string

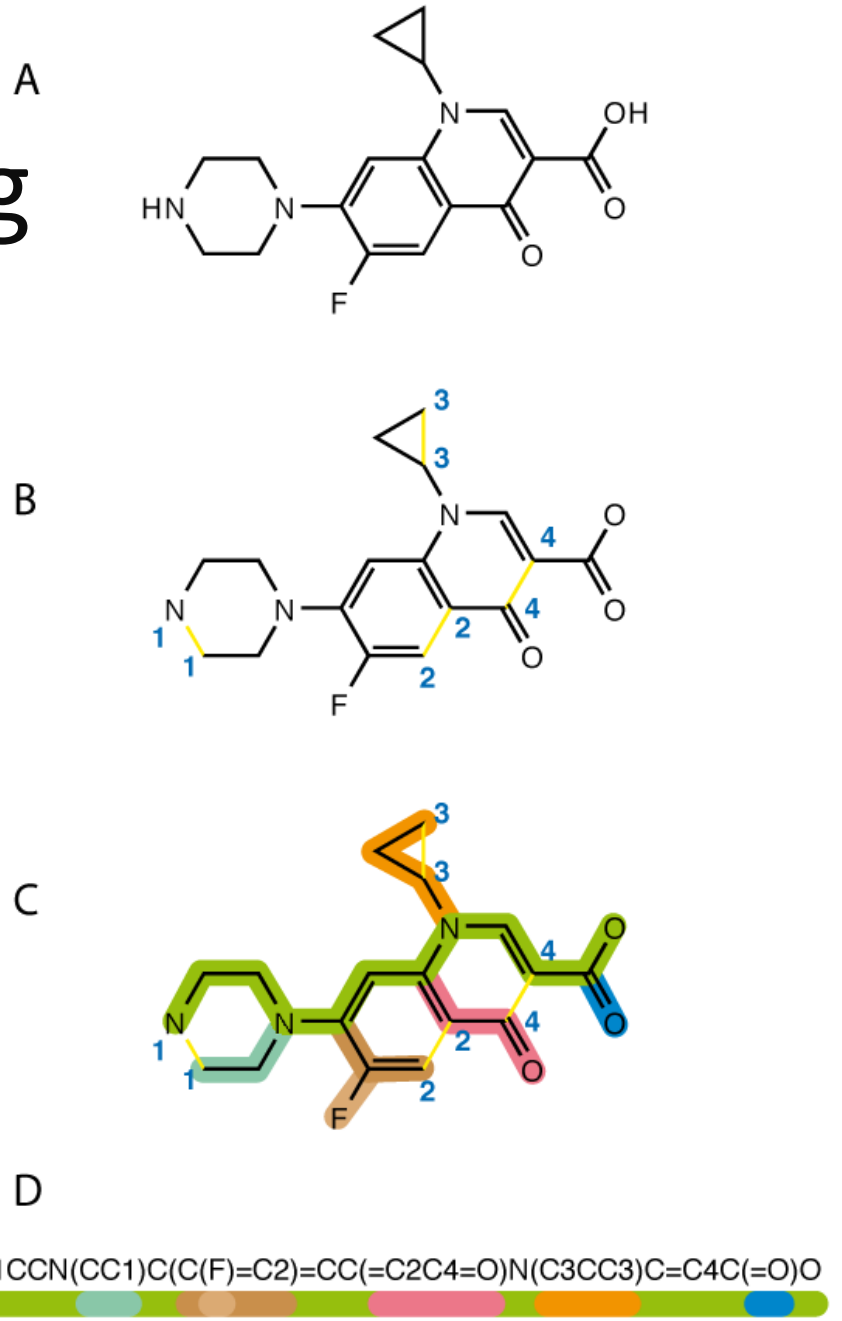
Using SMILES representation of drug, to convert a molecular graph into a string

- SMILES = Simplified Molecular-Input Line-Entry System

Then using sequence-to-sequence + VAE/GAN to model the continuous space that encodes/ decodes SMILES strings

- Allow easy optimization on the continuous space
- Problem: String → graphs is not unique!

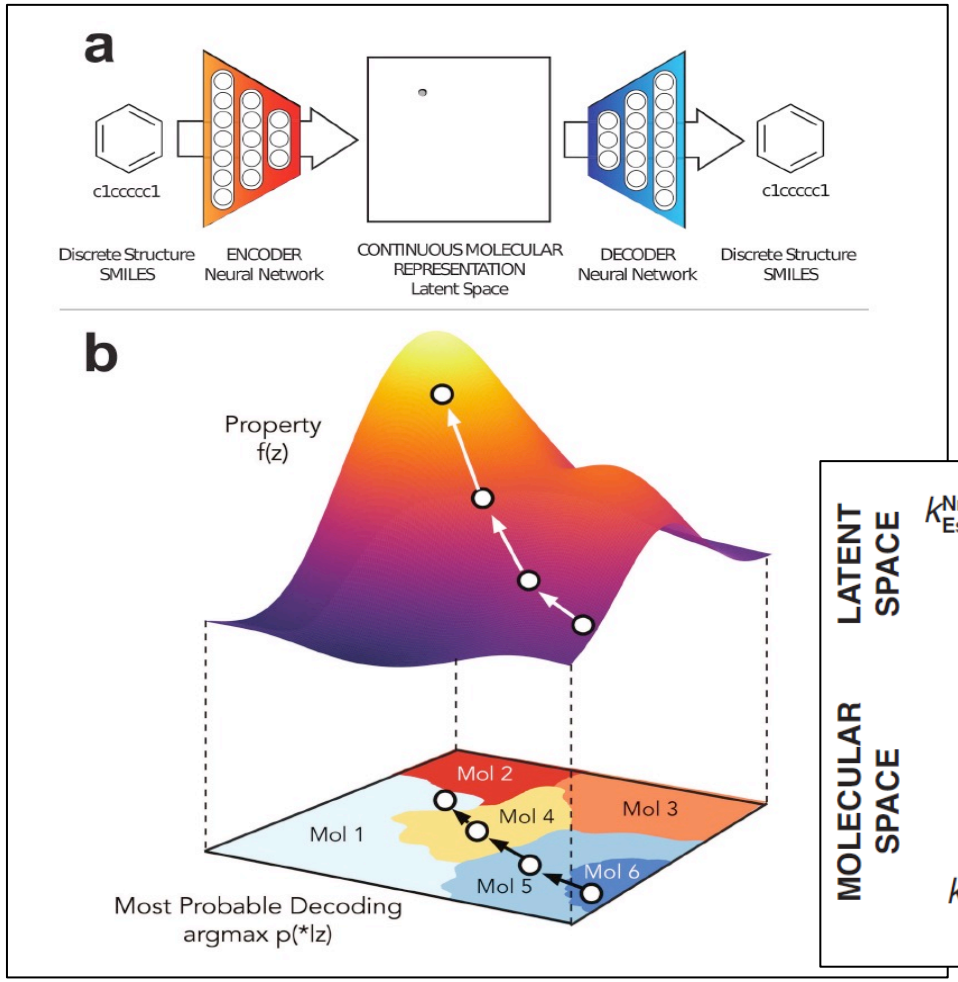
A better way is to encode/decode graph directly.



Source: wikipedia.org

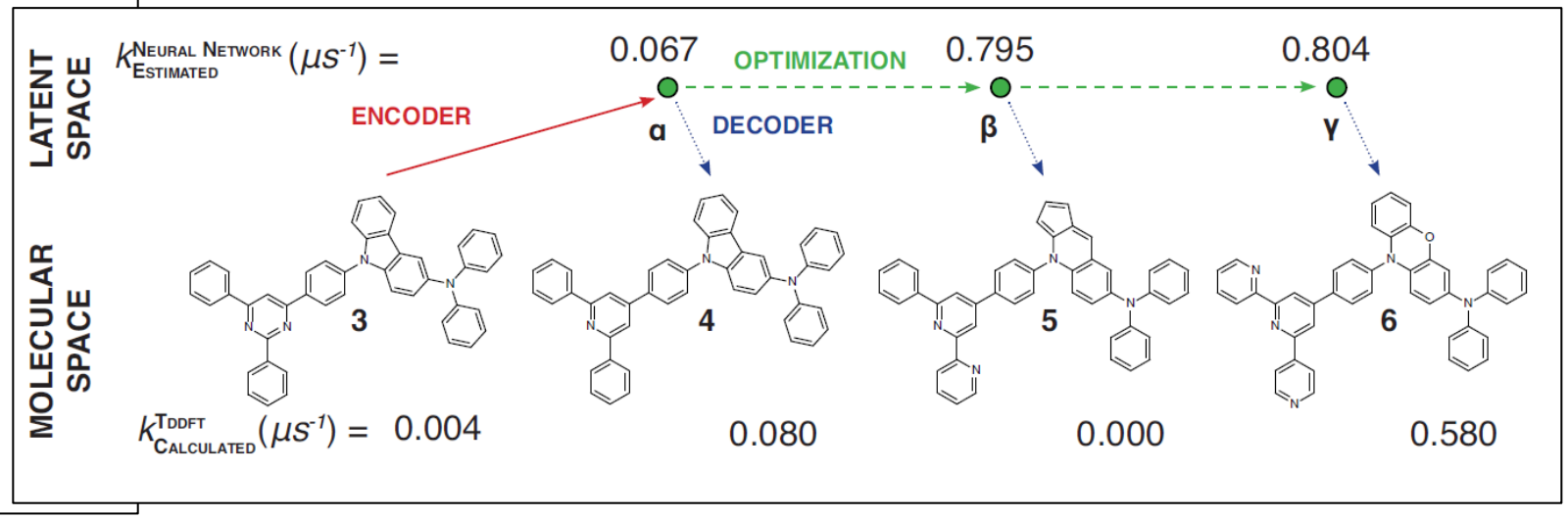
#REF: Gómez-Bombarelli, Rafael, et al. "Automatic chemical design using a data-driven continuous representation of molecules." *arXiv preprint arXiv:1610.02415* (2016).

VAE for drug space modelling



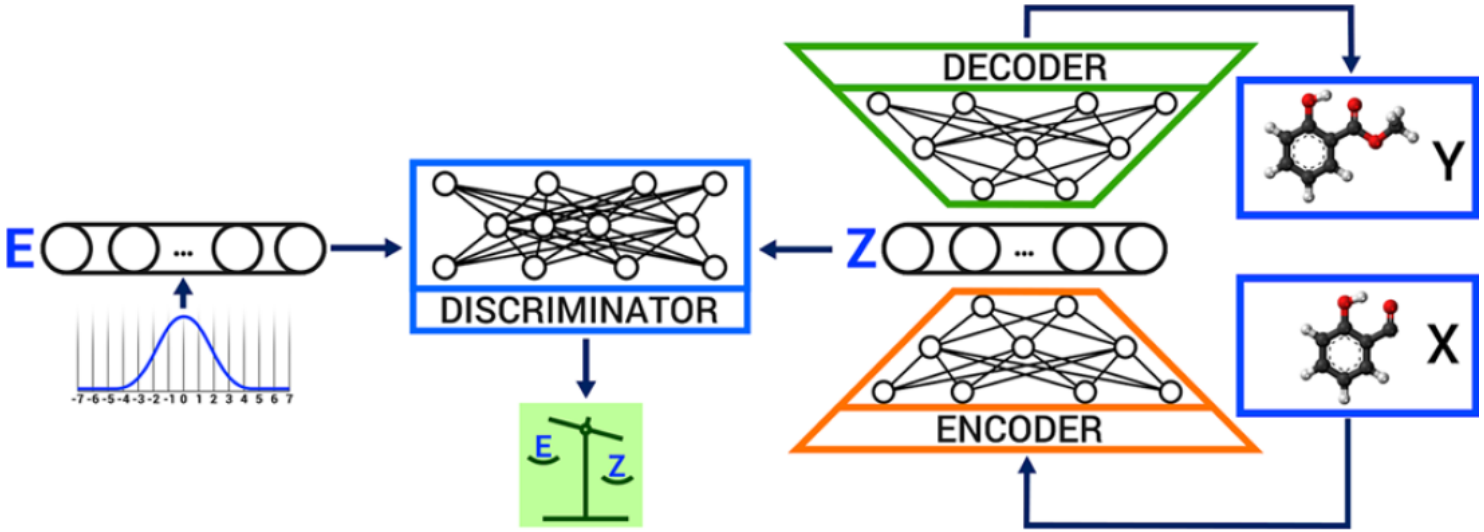
Uses VAE for sequence-to-sequence.

#REF: Bowman, Samuel R., et al. "Generating sentences from a continuous space." *arXiv preprint arXiv: 1511.06349* (2015).



Adversarial Autoencoders

Kadurin et al. Molecular pharmaceutics 2017



Input of the encoder : the fingerprint of a molecule

The decoder outputs the predicted fingerprint .

The generative model generates a vector E, which is then discriminated from the latent vector of the real molecule by the discriminator.

More refs

Segler, Marwin HS, et al. "Generating focussed molecule libraries for drug discovery with recurrent neural networks." *arXiv preprint arXiv: 1701.01329* (2017).

Kusner, Matt J., Brooks Paige, and José Miguel Hernández-Lobato. "Grammar Variational Autoencoder." *arXiv preprint arXiv:1703.01925* (2017).

Gupta, Anvita, et al. "Generative Recurrent Networks for De Novo Drug Design." *Molecular Informatics* (2017).

Living in the future: AI for health care

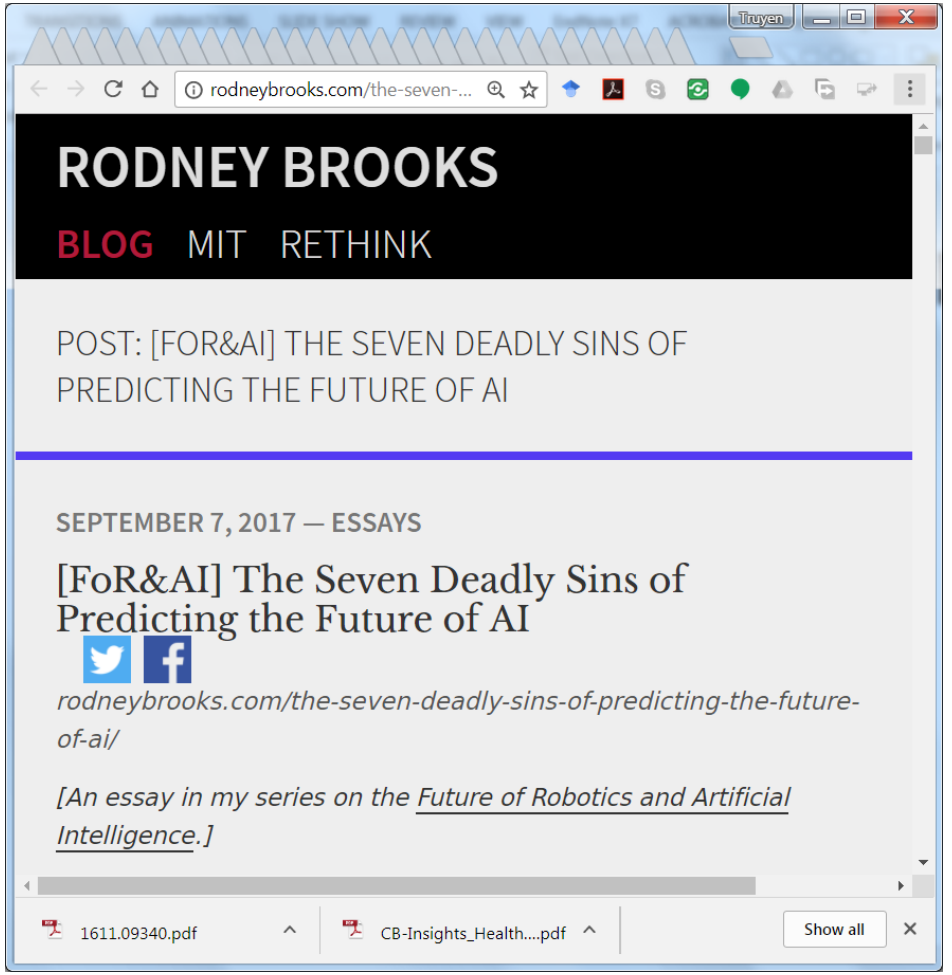
Some speculations (by me):

- https://letdataspeak.blogspot.com.au/2017/02/living-in-future-deep-learning-for.html

Bear in mind that anything beyond 5 years are nearly impossible to predict!

Kai-Fu Lee's vision:

- Wave 1: Internet data (→ PubMed, social media)
- **Wave 2: Business data (→EMR)**
- **Wave 3: Digitalize the physical world (→Drugs)**
- Wave 4: Full automation (→ Robot surgeons, GPs)



Toward personalize medicine

Will this patient response to that treatment?

Can we find the best treatment for a patient?

Which biomarkers predict the patient's response?

Sound familiar to Recommender Systems (patient = user, treatment = item)?

#REF: Talk by Chloé-Agathe Azencott titled "Machine learning for therapeutic research", 12/10/2017

Towards a dialog system → **Replace GP?**

Leveraging existing knowledge

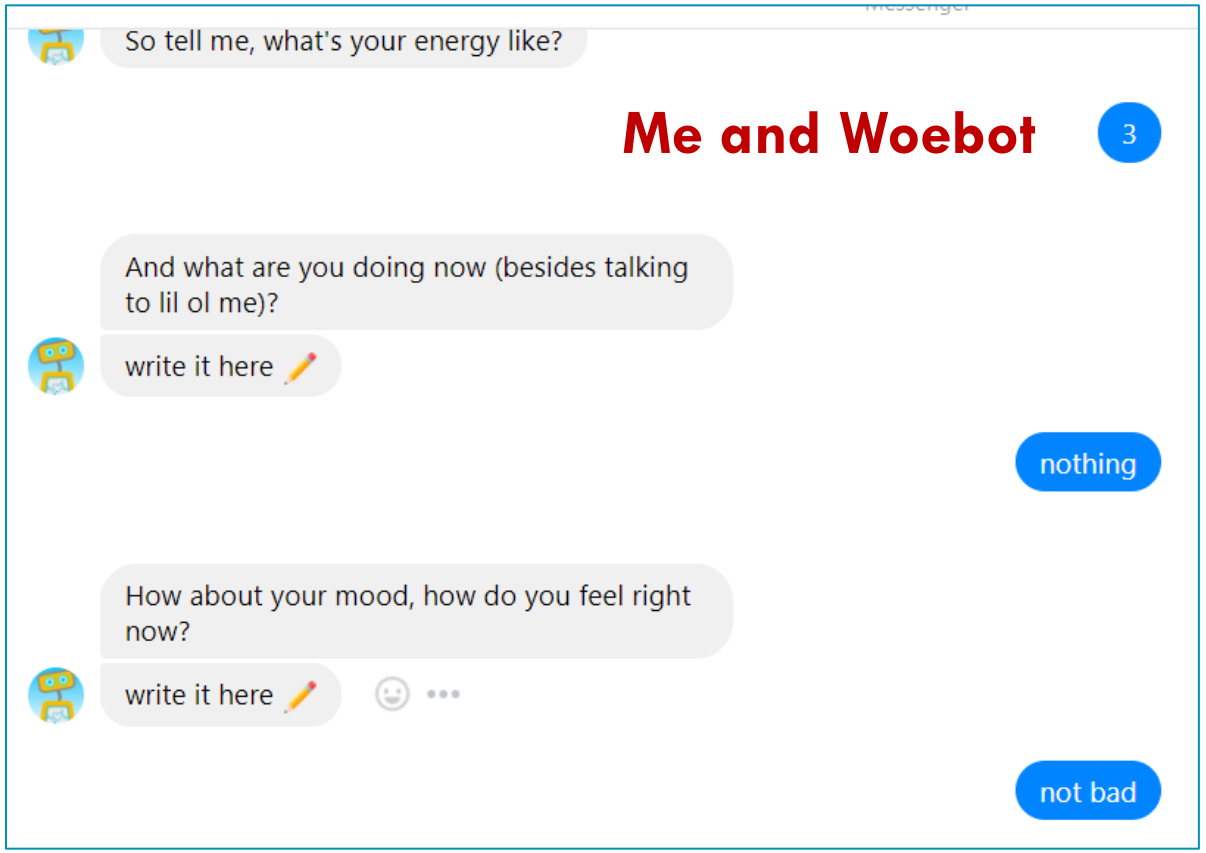
- Medical knowledge bases
- Medical texts
- Probably needs to build knowledge bases from text

Personalizing through EMRs

- Learn from hospitals data

Ask right questions → Finding answers from databases → Generating dialog

Never ending learning (NEL).



Other rooms for deep learning

-omics
- Gene expression
- Proteomics

Neuroscience
- Models for spike trains
- Deep learning for connectomics
- Neuroscience-inspired DL

Biomedical NLP
- Classical NLP
- Social media
- Knowledge graphs

Wearables
- Tracking the state of physical and mental health
- Lifestyle management & monitoring

Health Insurance
- Future illness/spending prediction
- Proactive prevention programs
- **WARNING: Working for insurance companies does raise ethical concerns!**

Nutrition
- Mobile phone vision → calories

Explainable AI
- Seeing through the black-box, e.g., visualization, motifs
- Explainable architectures that use biological mechanisms and medical ontologies
- Dual architecture: predictor & explainer

Thank you!

We're hiring

PhD & Postdocs

truyen.tran@deakin.edu.au

http://ahsanqawl.com/2015/10/qa/