Student-Teacher Curriculum Learning via Reinforcement Learning:

Predicting Hospital Inpatient Admission Location

Authors: Rasheed el-Bouri, Tingting Zhu, David Clifton

Clinical Authors: David Eyre, Peter Watkinson

lab

Computational
Health
Informatics

Reactive Environment:

- **emergency departments** (EDs) provide one of the greatest bottlenecks in the hospitalisation process – (more acute since pandemic!)

- Can we make this predictive?

So What's the problem if it takes a little time?

- Increased risk of adverse outcomes for patients.

- Some countries impose financial repercussions on hospitals for long patient waiting times.

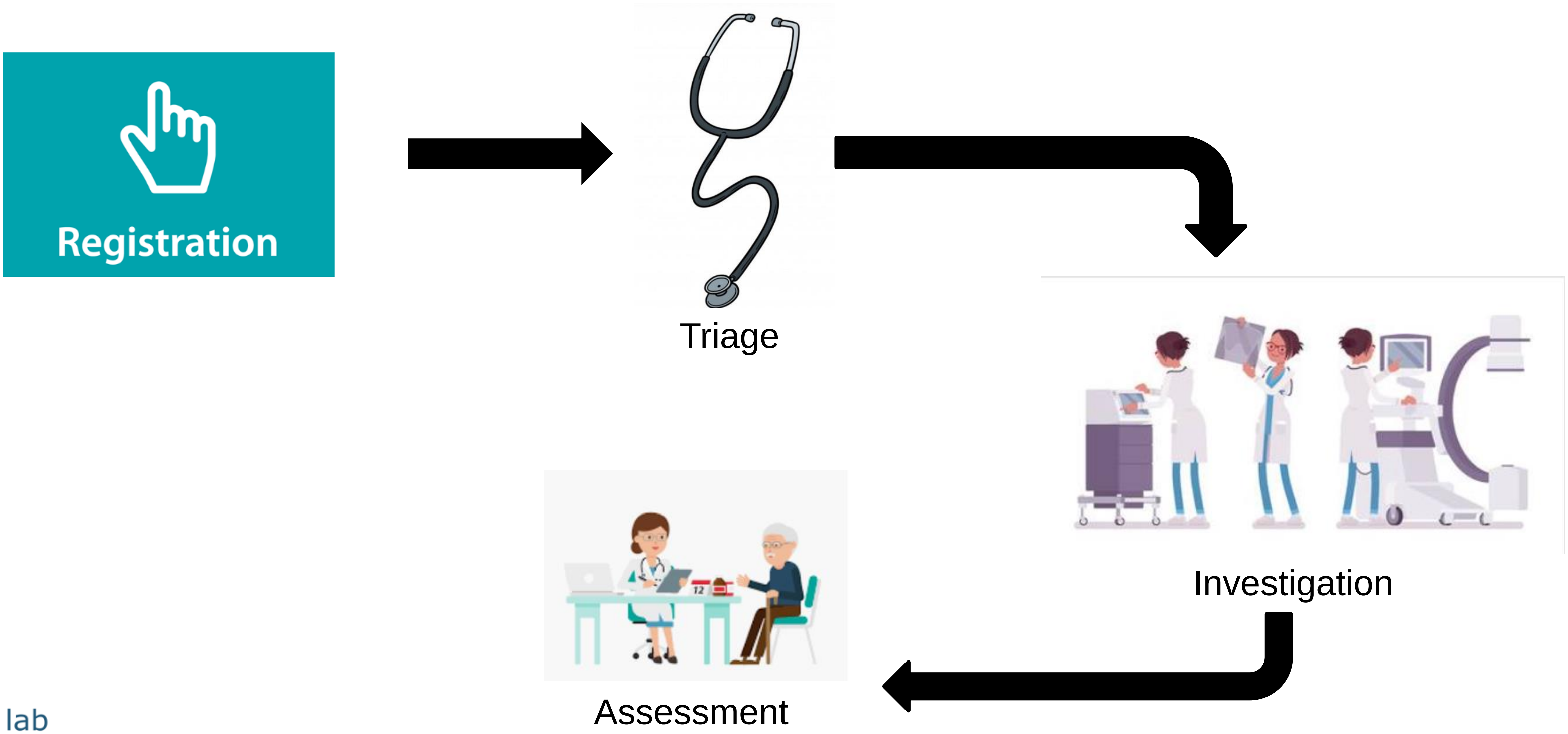
- Patients remaining in the ED still need to be cared for. Acts as a closed-loop and slows the entire process down even further.

Registration

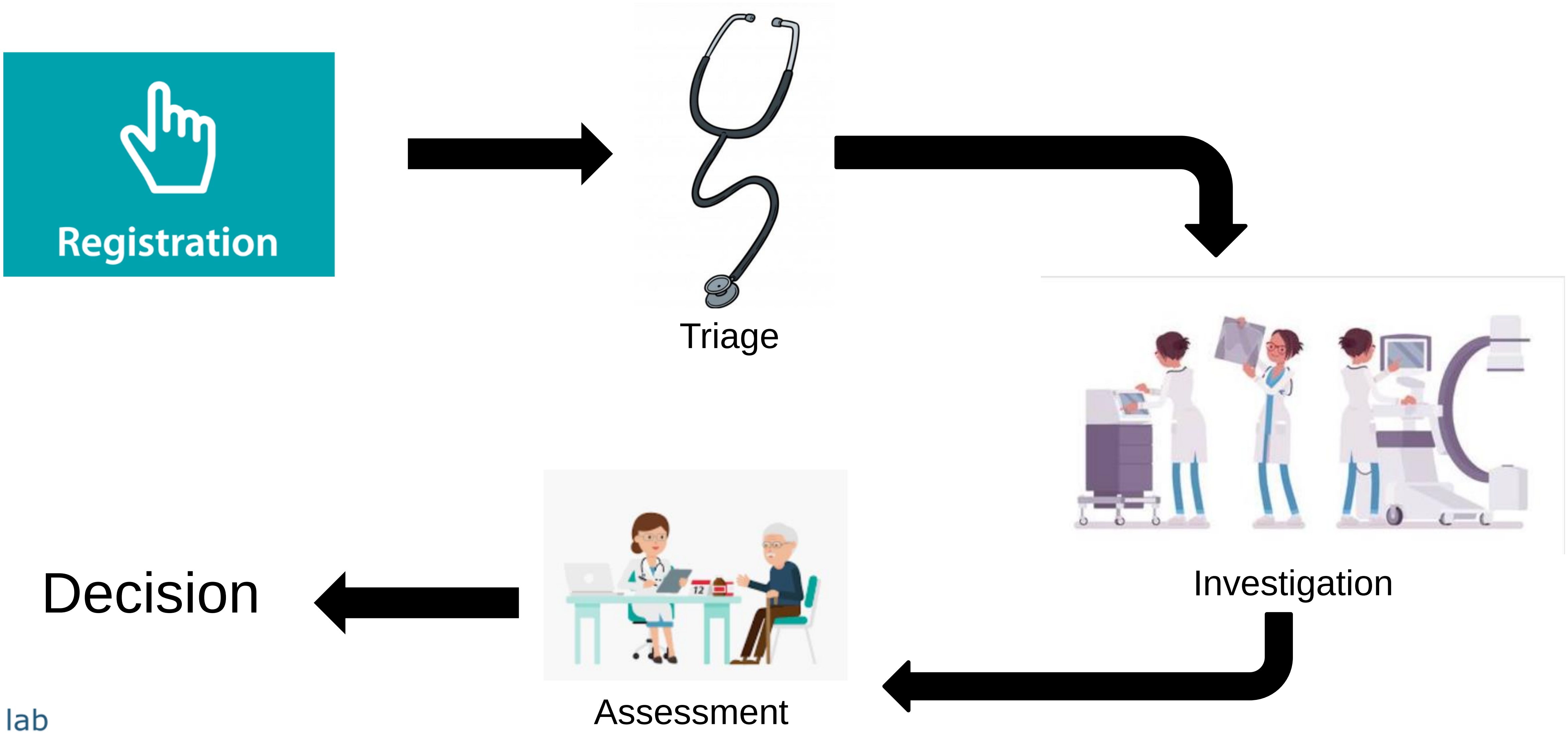
Triage

Triage

Investigation



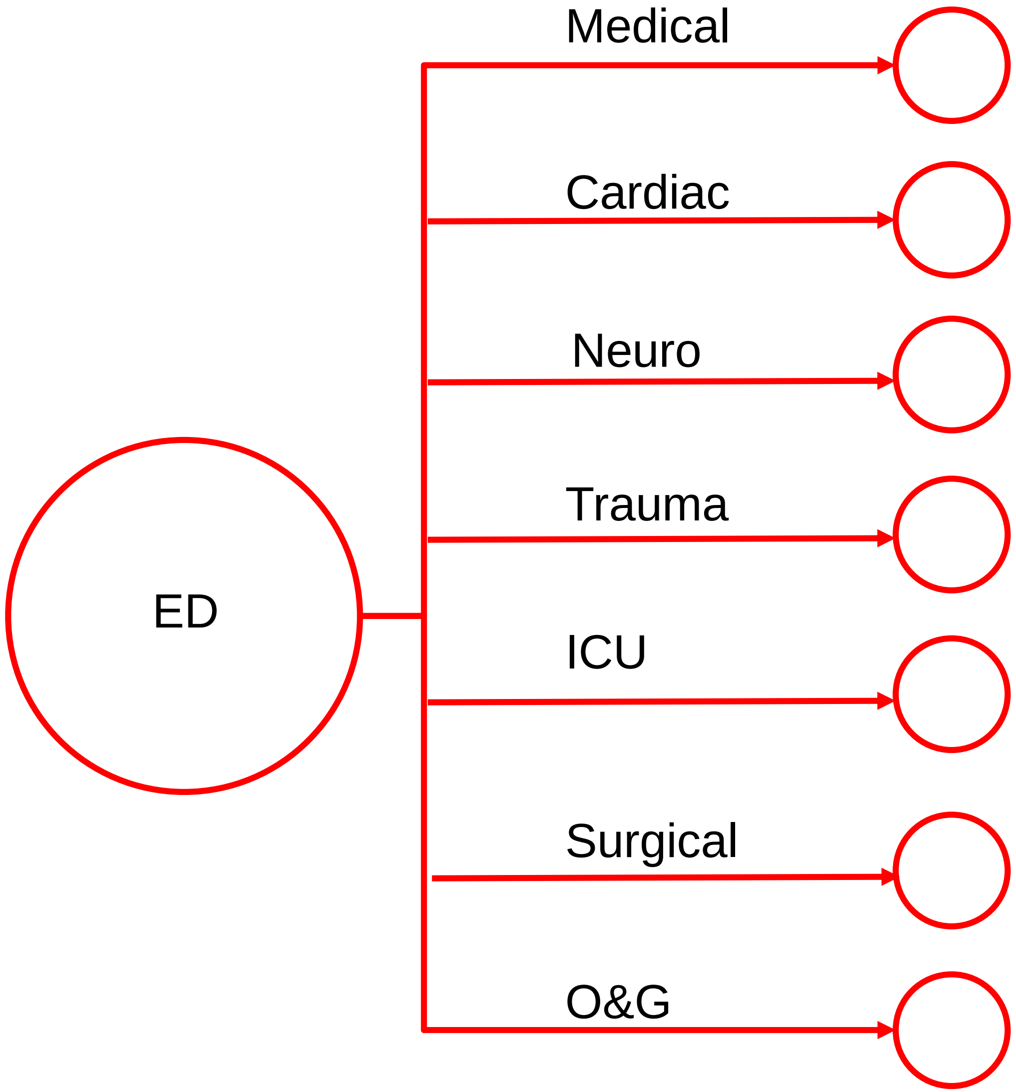
Problem outline



Problem outline

What do we Propose?

- Predict what type of ward a patient will be admitted to (seven classes).
- The type of ward is used so that any ward in the hospital with that functional capability can be considered
- An accurate answer as soon as the patient walks in is the most useful!

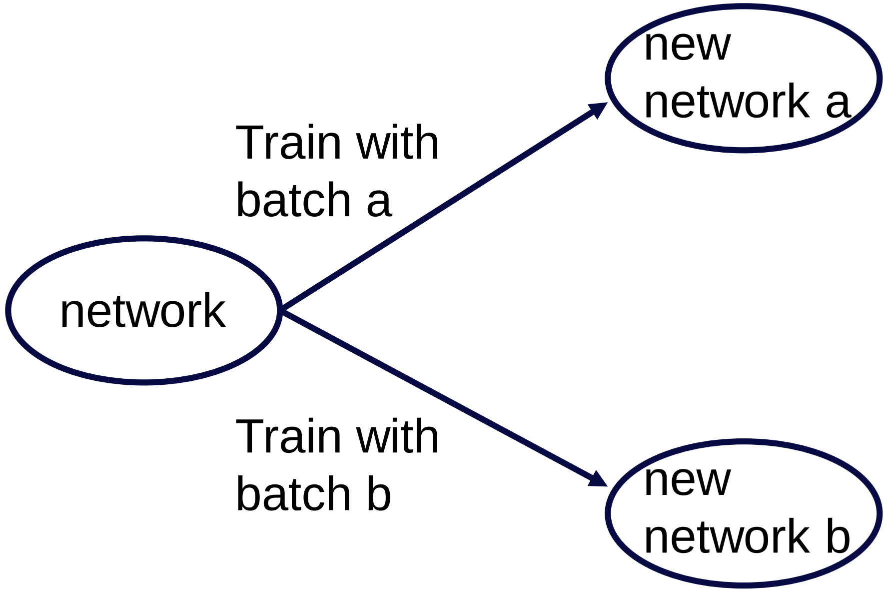


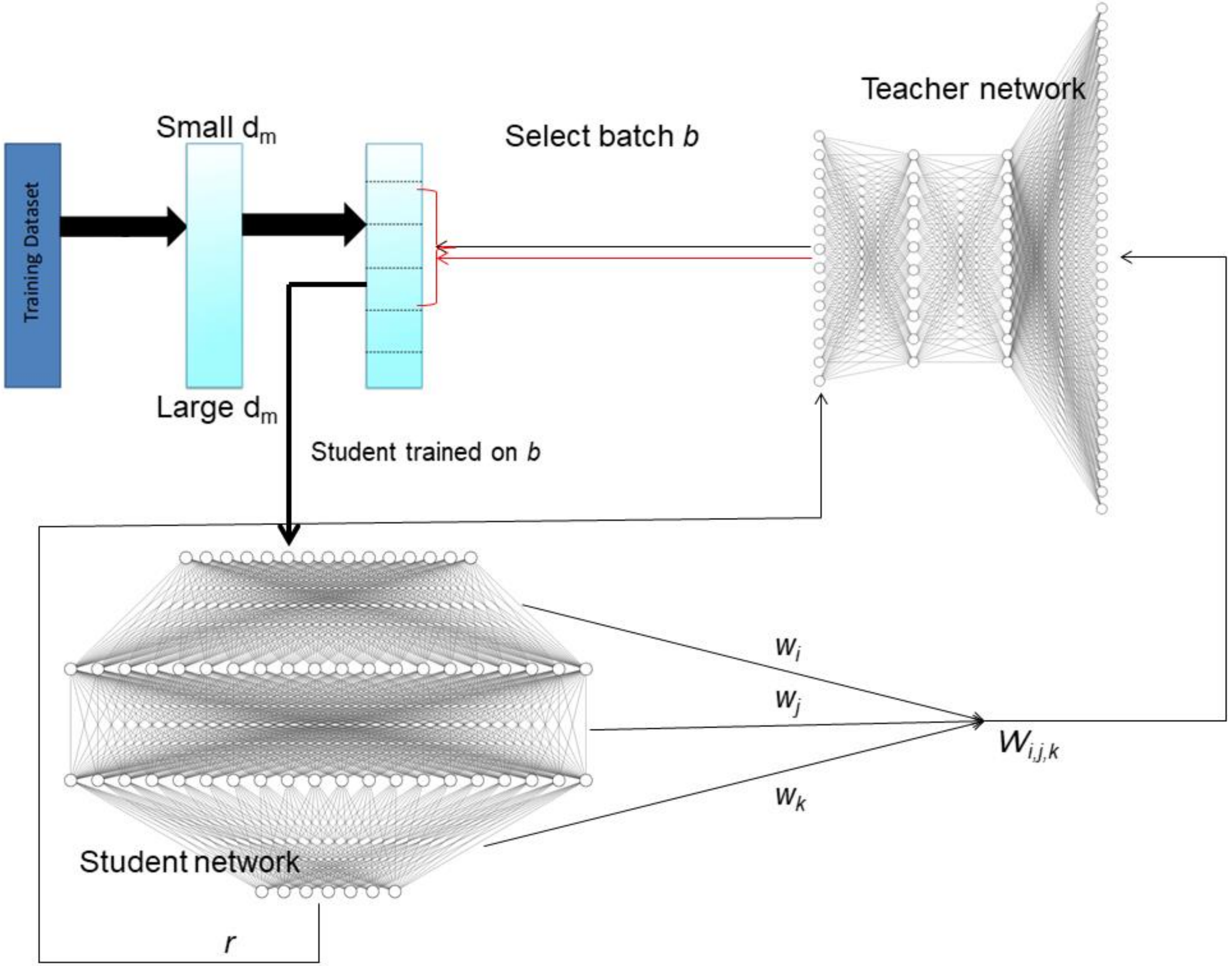
How can we do this?

- Curriculum learning has improved the performance of many algorithms that are trained using gradient descent.
- No real consensus on the best type of curriculum for a given problem
- Can we tailor-make a curriculum? Not just for a task but for a model too!

How can we do this?

- Training a neural network is Markovian

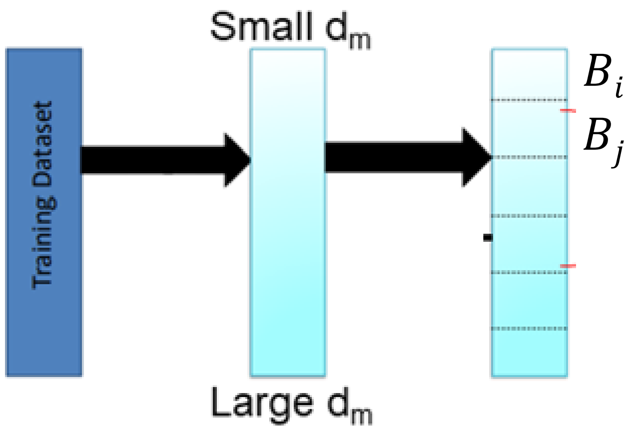




Student-teacher curriculum learning via reinforcement learning

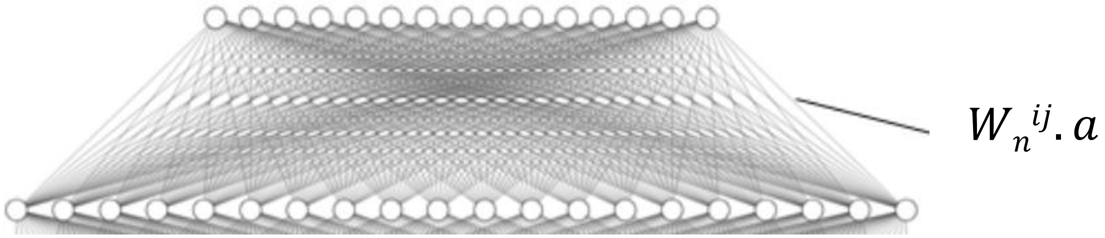
Generate the Curriculum

- Some similarity metric, H, that allows us to sort our training data set from most complex to least complex, such that $H[B_i] < H[Bj]$



- In this work we use cosine similarity as H for images and the Mahalanobis distance as H for categorical and numerical data

How do we calculate the student state?

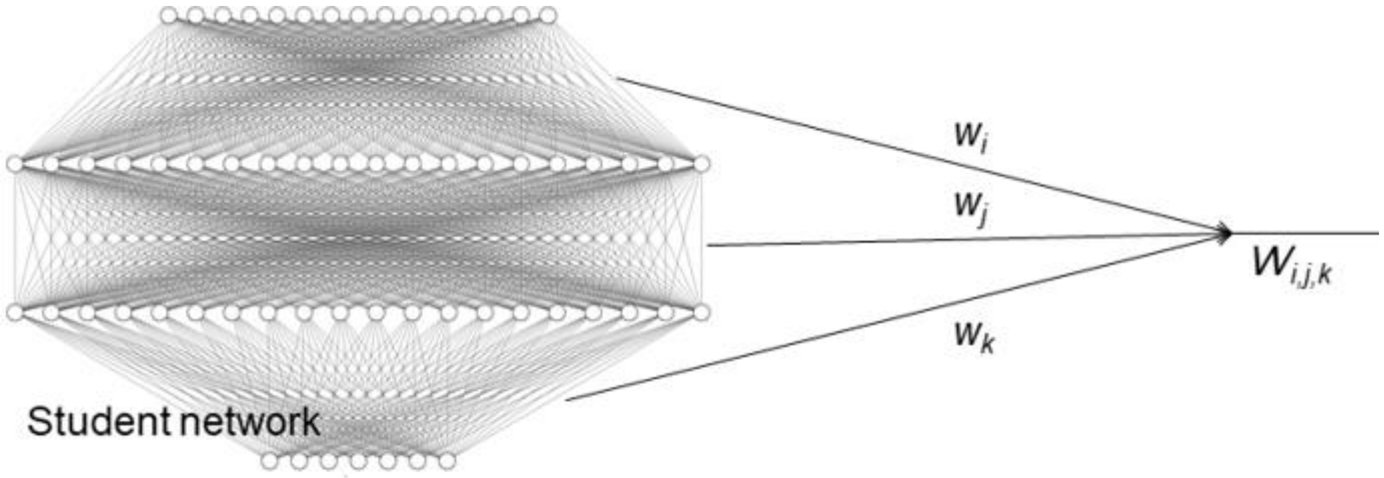


- Weights W^{ij} of size $M_i \times M_j$
- Reference vector of unique elements, a
- $|W_n^{ij}.a|$ and $\angle\,(W_n^{ij}.a)$

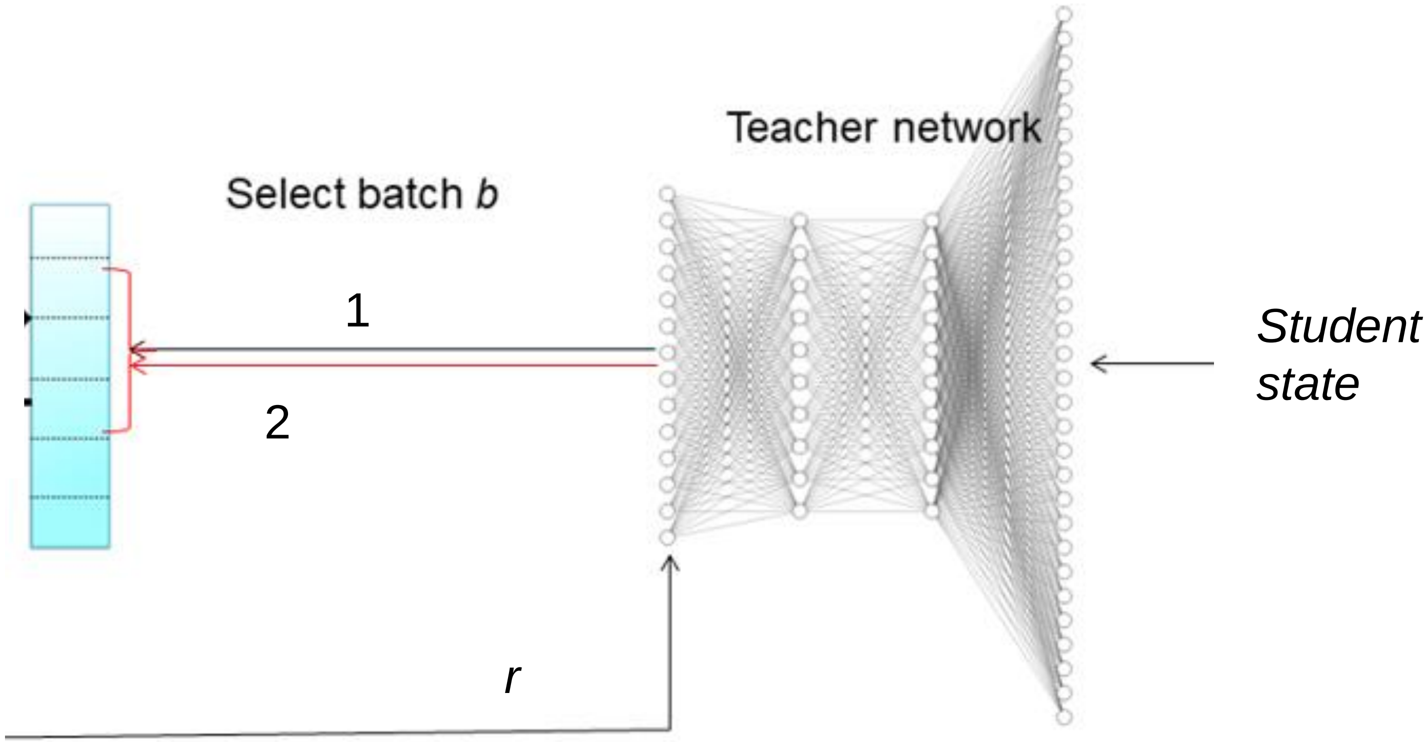
How do we calculate the student state? (2)

- Representation of layer i is $\mathbf{v}_i \in \mathbb{R}^{2M_i}$ and all the layers are concatenated together.

- The full representation of a student with k hidden layers is $\mathbf{v} \in \mathbb{R}^{2\left(\sum_0^k M_k\right)}$



How is the teacher trained? (2)



- DDPG or DQN
- $r = \nabla training * \nabla validation$
- The actor (teacher) has two outputs :

1. Curriculum index
2. Batchwidth

Results

Method	Ward Admission	MIMIC-III	CIFAR-10 Sample 1	CIFAR-10 Sample 2	CIFAR-10 Sample 3
	Acc (SD)	Acc (SD)	Acc (SD)	Acc (SD)	Acc (SD)
Batchwise	0.45 (0.01)	0.63 (0.01)	0.65 (0.02)	0.65 (0.02)	0.65 (0.02)
Curriculum	0.48 (0.02)	0.63 (0.02)	0.68 (0.01)	0.68 (0.01)	0.68 (0.01)
ST + curric	0.53 (0.03)	0.63 (0.02)	0.68 (0.02)	0.68 (0.02)	0.68 (0.02)
DeepFM	0.59 (0.01)	0.66 (0.01)	N/A	N/A	N/A
Deep+CrossNet	0.58 (0.02)	0.68 (0.02)	N/A	N/A	N/A
AutoInt	0.57 (0.02)	0.67 (0.01)	N/A	N/A	N/A
DenseNet*	N/A	N/A	0.96 (0.01)	0.96 (0.01)	0.96 (0.01)
GPipe*	N/A	N/A	**0.99** (0.01)	**0.99** (0.01)	**0.99** (0.01)
DQN student-teacher	0.58 (0.01)	0.66 (0.02)	0.86 (0.04)	0.88 (0.02)	0.86 (0.03)
DDPG student-teacher	**0.62** (0.02)	**0.70** (0.01)	0.90 (0.01)	0.90 (0.01)	0.89 (0.01)

Results

Method	Ward Admission Acc (SD)	MIMIC-III Acc (SD)	CIFAR-10 Sample 1 Acc (SD)	CIFAR-10 Sample 2 Acc (SD)	CIFAR-10 Sample 3 Acc (SD)
Batchwise	0.45 (0.01)	0.63 (0.01)	0.65 (0.02)	0.65 (0.02)	0.65 (0.02)
Curriculum	0.48 (0.02)	0.63 (0.02)	0.68 (0.01)	0.68 (0.01)	0.68 (0.01)
ST + curric	0.53 (0.03)	0.63 (0.02)	0.68 (0.02)	0.68 (0.02)	0.68 (0.02)
DeepFM	0.59 (0.01)	0.66 (0.01)	N/A	N/A	N/A
Deep+CrossNet	0.58 (0.02)	0.68 (0.02)	N/A	N/A	N/A
AutoInt	0.57 (0.02)	0.67 (0.01)	N/A	N/A	N/A
DenseNet*	N/A	N/A	0.96 (0.01)	0.96 (0.01)	0.96 (0.01)
GPipe*	N/A	N/A	**0.99** (0.01)	**0.99** (0.01)	**0.99** (0.01)
DQN student-teacher	0.58 (0.01)	0.66 (0.02)	0.86 (0.04)	0.88 (0.02)	0.86 (0.03)
DDPG student-teacher	**0.62** (0.02)	**0.70** (0.01)	0.90 (0.01)	0.90 (0.01)	0.89 (0.01)

Results

Method	Ward Admission Acc (SD)	MIMIC-III Acc (SD)	CIFAR-10 Sample 1 Acc (SD)	CIFAR-10 Sample 2 Acc (SD)	CIFAR-10 Sample 3 Acc (SD)
Batchwise	0.45 (0.01)	0.63 (0.01)	0.65 (0.02)	0.65 (0.02)	0.65 (0.02)
Curriculum	0.48 (0.02)	0.63 (0.02)	0.68 (0.01)	0.68 (0.01)	0.68 (0.01)
ST + curric	0.53 (0.03)	0.63 (0.02)	0.68 (0.02)	0.68 (0.02)	0.68 (0.02)
DeepFM	0.59 (0.01)	0.66 (0.01)	N/A	N/A	N/A
Deep+CrossNet	0.58 (0.02)	0.68 (0.02)	N/A	N/A	N/A
AutoInt	0.57 (0.02)	0.67 (0.01)	N/A	N/A	N/A
DenseNet*	N/A	N/A	0.96 (0.01)	0.96 (0.01)	0.96 (0.01)
GPipe*	N/A	N/A	**0.99** (0.01)	**0.99** (0.01)	**0.99** (0.01)
DQN student-teacher	0.58 (0.01)	0.66 (0.02)	0.86 (0.04)	0.88 (0.02)	0.86 (0.03)
DDPG student-teacher	**0.62** (0.02)	**0.70** (0.01)	0.90 (0.01)	0.90 (0.01)	0.89 (0.01)

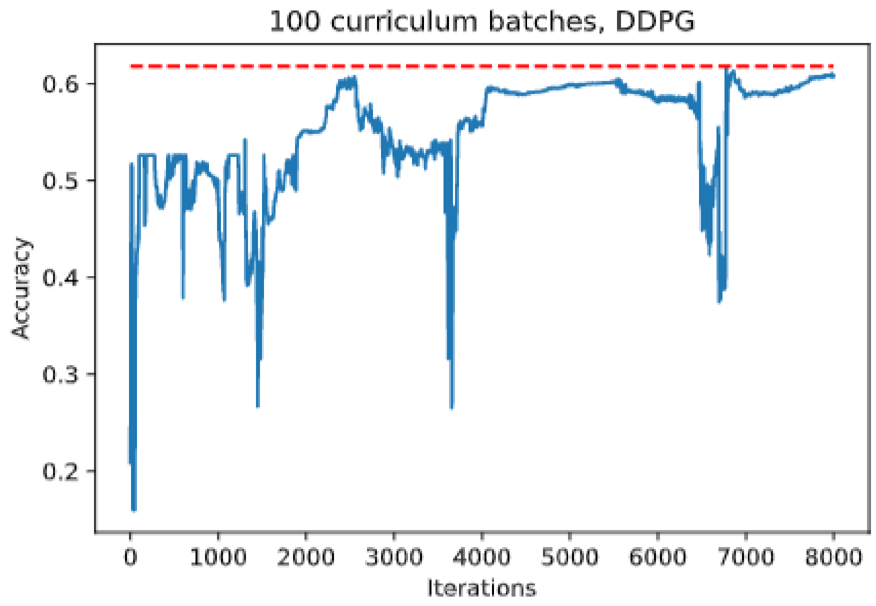
Results

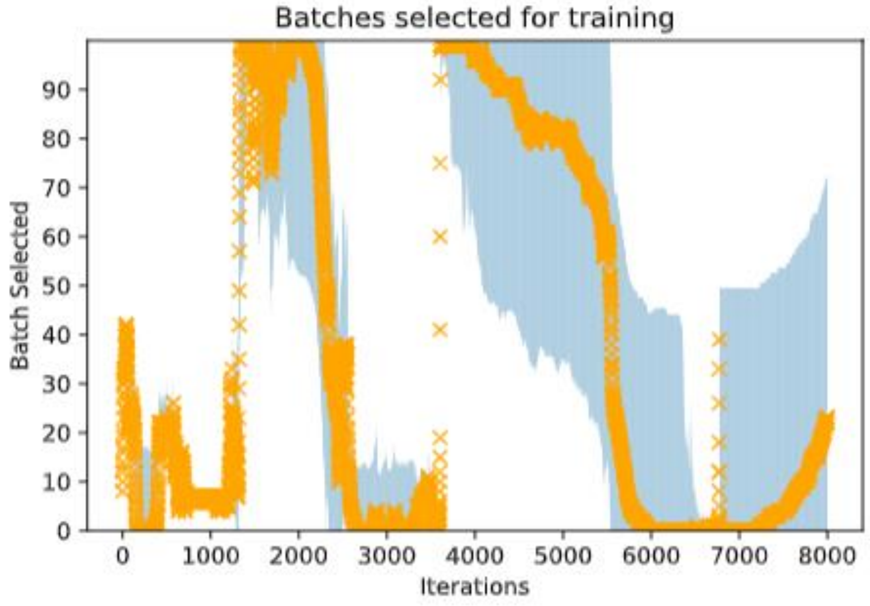
Method	Ward Admission Acc (SD)	MIMIC-III Acc (SD)	CIFAR-10 Sample 1 Acc (SD)	CIFAR-10 Sample 2 Acc (SD)	CIFAR-10 Sample 3 Acc (SD)
Batchwise	0.45 (0.01)	0.63 (0.01)	0.65 (0.02)	0.65 (0.02)	0.65 (0.02)
Curriculum	0.48 (0.02)	0.63 (0.02)	0.68 (0.01)	0.68 (0.01)	0.68 (0.01)
ST + curric	0.53 (0.03)	0.63 (0.02)	0.68 (0.02)	0.68 (0.02)	0.68 (0.02)
DeepFM	0.59 (0.01)	0.66 (0.01)	N/A	N/A	N/A
Deep+CrossNet	0.58 (0.02)	0.68 (0.02)	N/A	N/A	N/A
AutoInt	0.57 (0.02)	0.67 (0.01)	N/A	N/A	N/A
DenseNet*	N/A	N/A	0.96 (0.01)	0.96 (0.01)	0.96 (0.01)
GPipe*	N/A	N/A	**0.99** (0.01)	**0.99** (0.01)	**0.99** (0.01)
DQN student-teacher	0.58 (0.01)	0.66 (0.02)	0.86 (0.04)	0.88 (0.02)	0.86 (0.03)
DDPG student-teacher	**0.62** (0.02)	**0.70** (0.01)	0.90 (0.01)	0.90 (0.01)	0.89 (0.01)

Learned Curricula

- Teacher learns how to degrade performance in order to start again and achieve a better performance after training again

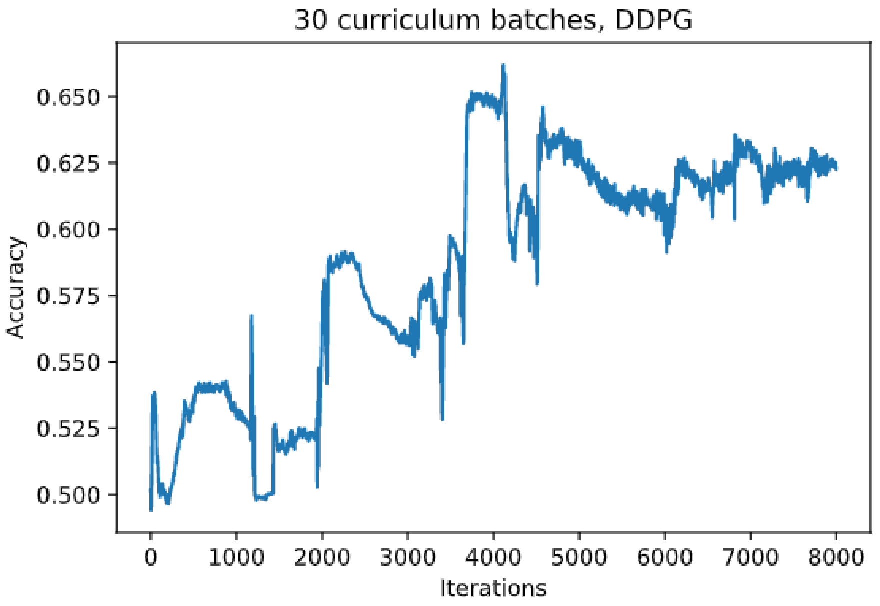
- Bottom plot: orange is first output index of teacher (index in curriculum), blue is second (width)

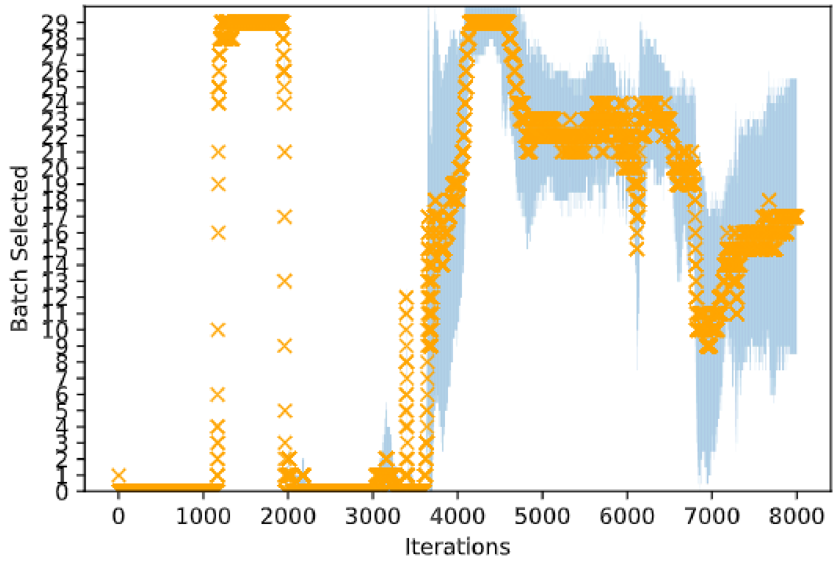




Capacity for Transfer Learning

- Plot shows performance on MIMIC-III mortality prediction task
- The teacher uses the same strategy as on the ward admission dataset and achieves a strong performance for the student
- Some metric of task similarity will allow teacher transfer for training





More Experiments in the Paper

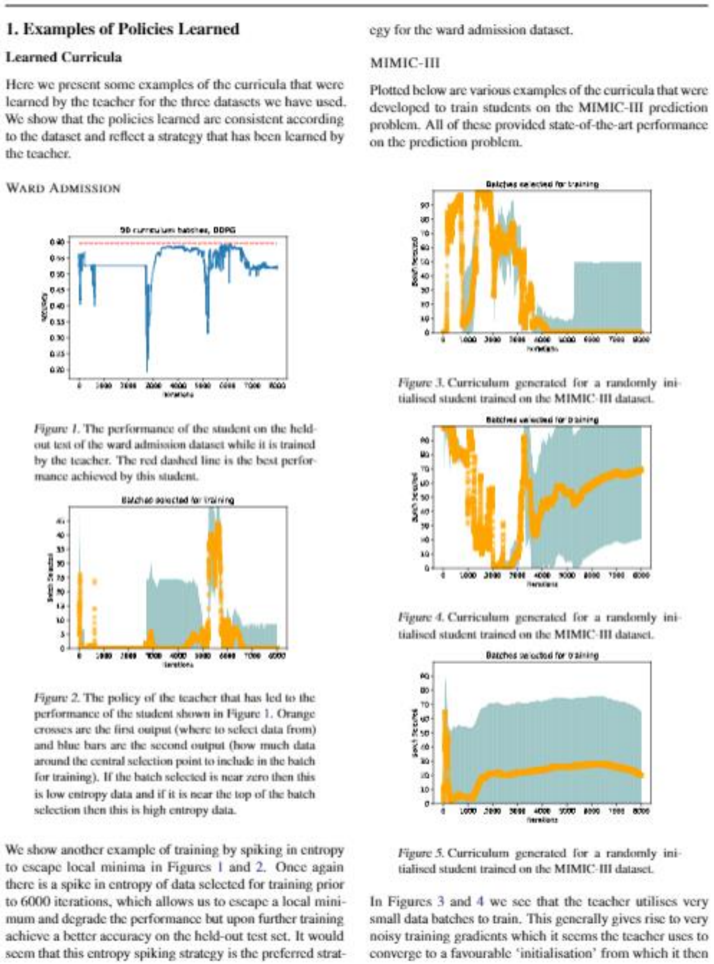
Discussion of curricula learned for all tasks

Policy Transfer between Tasks

Constrained teacher

Convergence of teacher selection

Policy calibration



1. Examples of Policies Learned

Learned Curricula

Here we present some examples of the curricula that were learned by the teacher for the three datasets we have used. We show that the policies learned are consistent according to the dataset and reflect a strategy that has been learned by the teacher.

WARD ADMISSION

Figure 1. The performance of the student on the held-out test of the ward admission dataset while it is trained by the teacher. The red dashed line is the best performance achieved by this student.

Figure 2. The policy of the teacher that has led to the performance of the student shown in Figure 1. Orange crosses are the first output (where to select data from) and blue bars are the second output (how much data around the central selection point to include in the batch for training). If the batch selected is near zero then this is low entropy data and if it is near the top of the batch selection then this is high entropy data.

We show another example of training by spiking in entropy to escape local minima in Figures 1 and 2. Once again there is a spike in entropy of data selected for training prior to 6000 iterations, which allows us to escape a local minimum and degrade the performance but upon further training achieve a better accuracy on the held-out test set. It would seem that this entropy spiking strategy is the preferred strategy for the ward admission dataset.

MIMIC-III

Plotted below are various examples of the curricula that were developed to train students on the MIMIC-III prediction problem. All of these provided state-of-the-art performance on the prediction problem.

Figure 3. Curriculum generated for a randomly initialised student trained on the MIMIC-III dataset.

Figure 4. Curriculum generated for a randomly initialised student trained on the MIMIC-III dataset.

Figure 5. Curriculum generated for a randomly initialised student trained on the MIMIC-III dataset.

In Figures 3 and 4 we see that the teacher utilises very small data batches to train. This generally gives rise to very noisy training gradients which it seems the teacher uses to converge to a favourable 'initialisation' from which it then

Thanks for listening!

Feel free to contact me at rasheed.el-bouri@eng.ox.ac.uk with any questions or ideas