Towards a Green Intelligence Applied to Health Care and Well-Being

John A. GARCÍA. H.

Frédéric PRECIOSO, Pascal STACCINI, Michel RIVEILL

Laboratoire d'Informatique, Signaux et Systèmes de Sophia Antipolis - I3S

Main goal

1) Exploit medico-administrative data collected by hospitals (EHR)

2) Do it on computer equipment

- **at a lower cost to be able to be installed in the hospitals and exploit the data directly on their place of production (preservation of privacy)**

- ***at low energy footprint***

3) Propose an entire workflow based on artificial neural networks approach

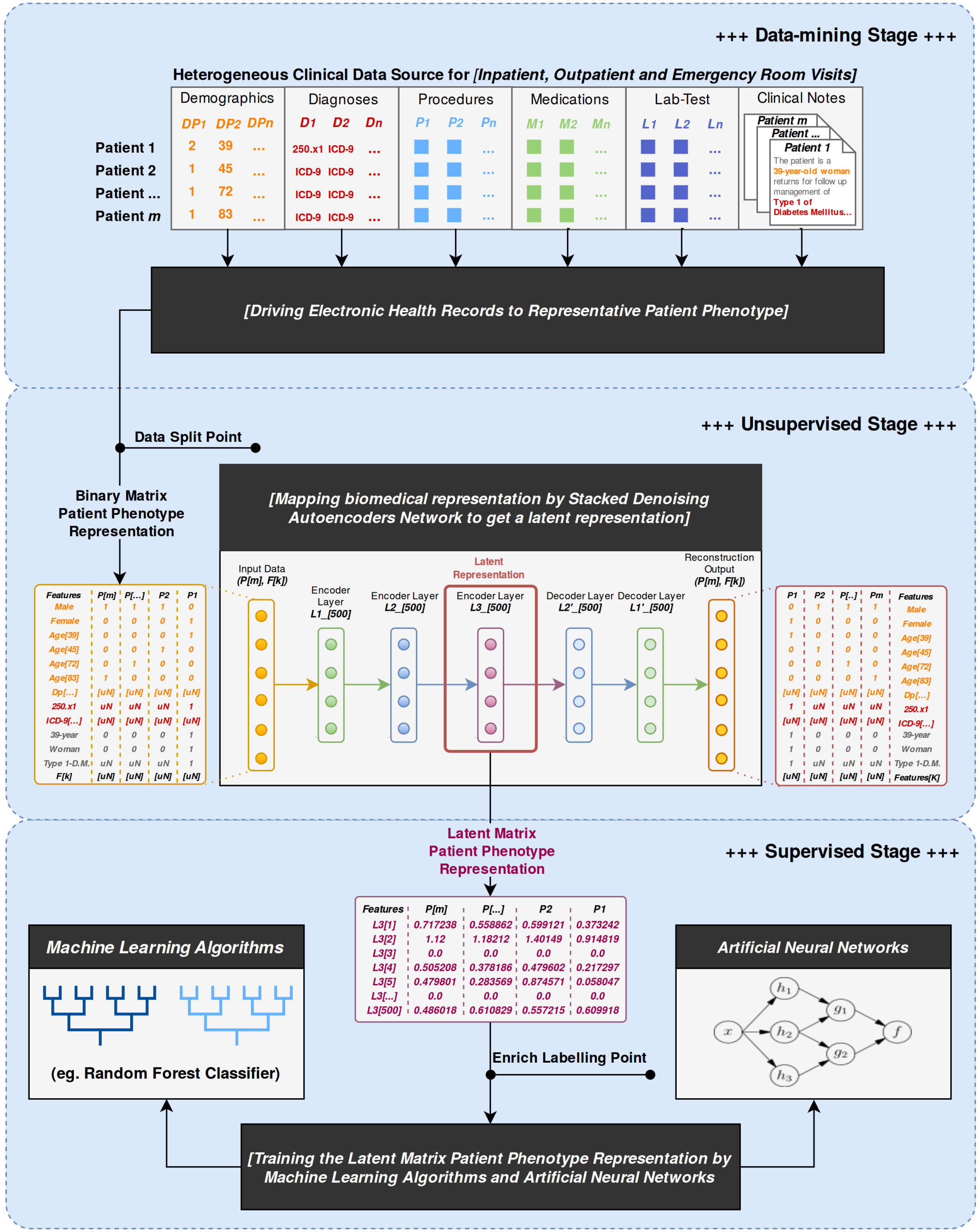
1. Source: **Doctor AI: Predicting Clinical Events via Recurrent Neural Networks**. By Edward Choi et al. Machine Learning for Healthcare 2016.
2. Source: **ICLR 2018: International Conference on Learning Representations.**
3. **Source: Exascale Computing Study: Technology Challenges in Achieving Exascale Systems.** Peter Kogge, Editor & Study Lead, Darpa 2008.

dIAgnoseNet: Framework to Build a Full Deep Neural Networks Workflow

Stage 1.
Data-mining stage &
Feature extraction:
Driving EHRs to build a binary phenotype representation.

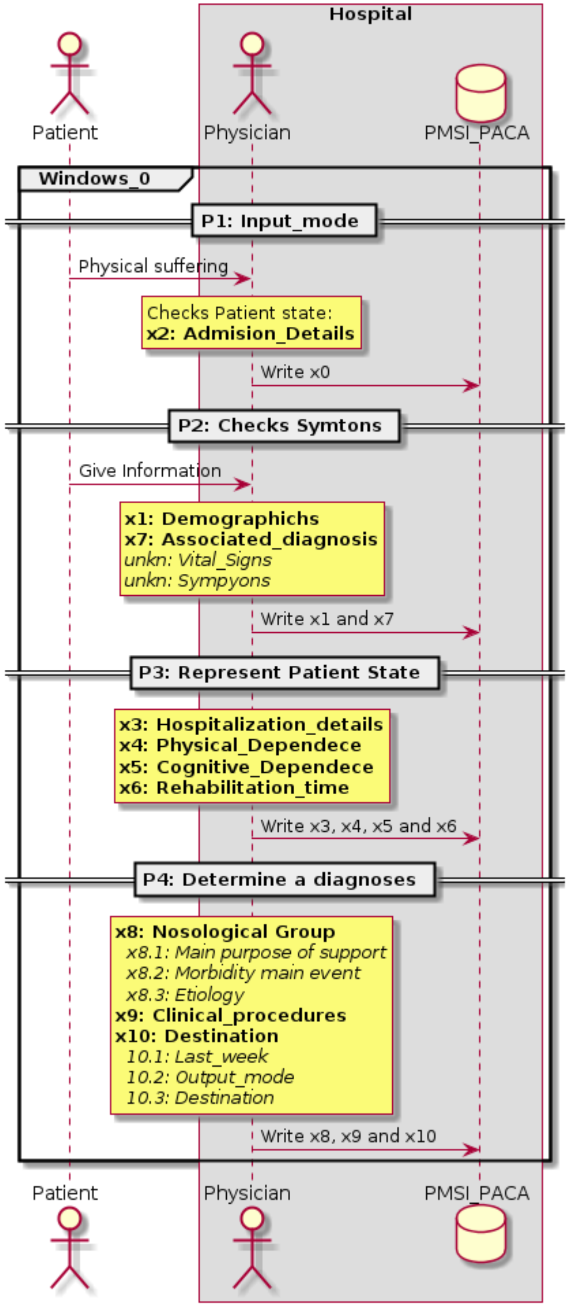
Eventualy Stage 2.
Unsupervised stage:
Mapping the Binary Patient PR.
to get a new space call Deep Patient
(or Latent Representation)
Using Stacked Denoising Autoencoders.

Stage 3.
Supervised stage:
Labelling Medical Target and
training the *Latent Representation*
by ML algorithms and ANN for classification and prediction of patient's disease.



1. Adapted from: **Deep Patient: An Unsupervised Representation to Predict the Future of Patients from the Electronic Health Records**. By Riccardo Miotto et al. SCIENTIFIC REPORTS, 2016.

Case of Study: Predict the Medical Future of Patients from EHRs



EHR is a iterative procedure

- **1 entry by week**
- **Medico-administrative information** used by government to give money back to hopital

Medical Targets at ICU in PACA

MT-1: Predict the ***'Major Clinical Category'*** from demographic, admission detail, diagnosis information

MT-2: Predict the ***'Clinical Procedures'*** from demographic, admission detail, diagnosis information and primary morbidity

MT-3: Predict the ***'Patient Destination'*** (home, transfer or death) and the ***'length of hospitalization stay'*** from demographic, admission detail, diagnosis information, primary morbidity and clinical procedures.

Stage 1. dIAgnoseNet Data-mining

Dynamic API

To Drive the Binary Patient Phenotype Representation (BPPR)

By Feature Extraction From Electronic Health Records.

+++ Data-mining & Feature Extraction +++

(Nominal Variables) X2 | (Ratio Variables) X3 | (Ordinal Variables) X4–X5 | (Ratio Variables) X6 | (ICD-10 Codes) X7–X8

	X1: Demographics			X2: Admission_details			X3: Hospitalisation_details			X4: Physical_dependence			X5: Cognitive_dependence			X6: Rehabilitation_time			X7: Related_diagnoses			X8:
	sex	...	age_group	input_mode	...	Input_source	days_hospitalized	...	week_sequence	feeding	...	displacement	communication	...	comportment	mechanical_rehab	...	kidneys_rehab	DA_1	DA_2	...	etiology
Patient 1	2	...	61	7	...	1	7	...	1	4	...	4	4	...	4	30	...	0	Z431	Z501	...	I619
Patient 2	2	...	62	7	...	1	7	...	52	4	...	4	4	...	4	25	...	0	Z431	Z501	...	I619
Patient ...	2	...	69	7	...	1	7	...	14	1	...	2	1	...	3	0	...	30	J459	F322	...	C20
Patient *m*	1	...	95	6	...	1	7	...	3	1	...	4	1	...	1	0	...	0	C259	0	...	0

Driving Electronic Health Records to Build Representative Binary of Patient Phenotype

i. diagnosenet_cdajson | ii. diagnosenet_featurescomposition | iii. diagnosenet_vocabularycomposition | iv. diagnosenet_dtm | v. diagnosenet_labelcomposition

Binary Patient Phenotype Representation

	[1: male]	[2: female]	...	60-74	+74	...	[6: Mutation]	[7: Normal_transfer]	[1: From a MCO unit]	...	[1: Complete_Independence]	[4: Total_Assistance]	...	Z431	Z501	J459	F322	C259	...	I619	C20	'0'	...
Patient 1	0	1	...	1	0	...	0	1	1	...	0	1	...	1	1	0	0	0	...	1	0	0	...
Patient 2	0	1	...	1	0	...	0	1	1	...	0	1	...	1	1	0	0	0	...	1	0	0	...
Patient ...	0	1	...	1	0	...	0	1	1	...	1	0	...	0	0	1	1	0	...	0	1	0	...
Patient *m*	1	0	...	0	1	...	1	0	1	...	1	0	...	0	0	0	0	1	...	0	0	1	...

BPPR extraction from integrated database

Dynamic API
To Drive the Binary Patient Phenotype Representation (BPPR)
By Feature Extraction From Electronic Health Records.

- Selection of a feature set → { X_i, ⋯ }
- Selection of an entre set (12,7 % - 100 K entries)

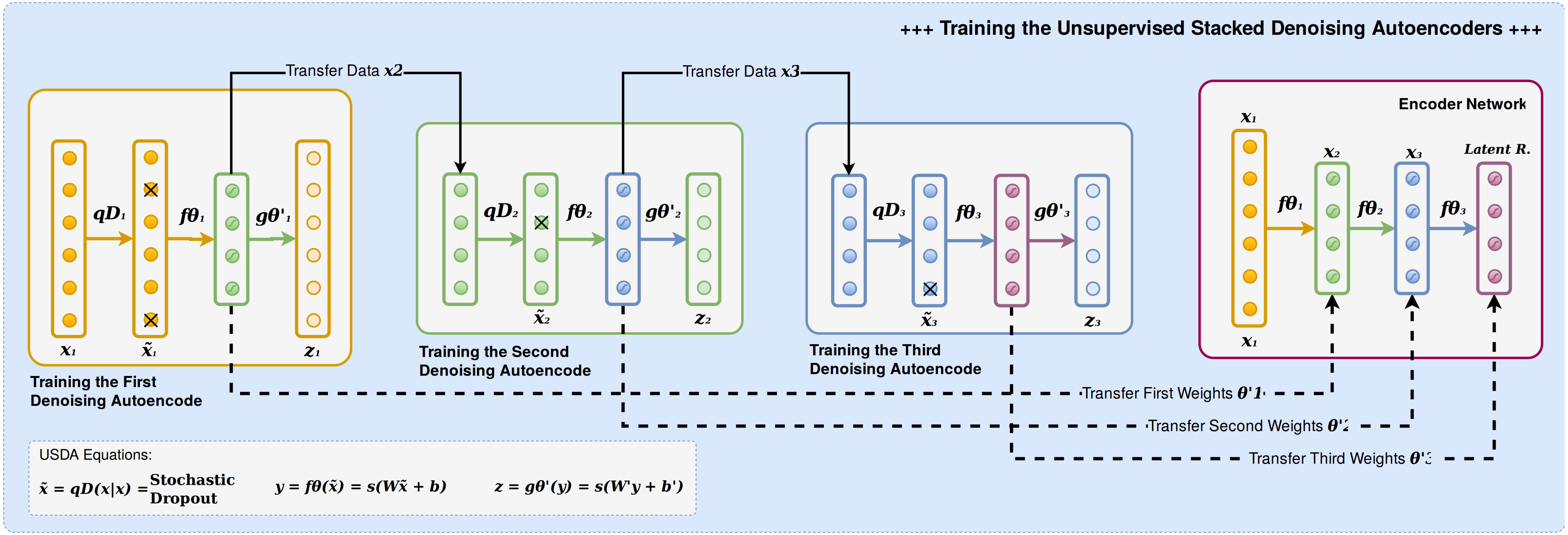
EHR Common Version	Features Composition	Binary Patient Phenotype Representation	BPPR (Features)	Patients (Samples)	Disk (size)	Exe. Time server (mins)
2007-2008	[X1, X2, X3, X4, X5, X6, X7]	BPPR 1	11094	99999	3.2 GB	1.86
	[X1, X2, X3, X4, X5, X6, X7, X8.3]	BPPR 2	14515	99999	4.1 GB	2.63
Without: X6 (Rehabilitation Time)	[X1, X2, X3, X4, X5, X7]	BPPR 3	8041	99999	2.3 GB	1.48
	[X1, X2, X3, X4, X5, X7, X8.3]	BPPR 4	11462	99999	3.3 GB	2.18

Data Split Point:

Patients (12.7 %)	No. (Samples) 99999	Mini Batch (Files) 11
Data Training	84999	9
Data Validation	4950	1
Data Test	10050	1

Stage 2. dIAgnoseNet Unsupervised Embedding

For Mapping the BPPR Through Unsupervised Stacked Denoising Autoencoder to get a New Encoded Space of Patient Phenotype (Latent Representation)



1. Adapted from **Stacked Denoising Autoencoders: Learning Useful Representations in a Deep Network with a Local Denoising Criterion.** By Pascal Vincent et al. 2010

dIAgnoseNet Unsupervised Embedding: Experiments to Predict the First Medical Target

For Mapping the BPPR Through Unsupervised Stacked Denoising Autoencoder to get a New Encoded Space of Patient Phenotype (Latent Representation)

EHR Common Version	Features Composition	BPPR (Features)	Patients (Samples)	BPPR Mini Batch (All files)	BPPR Mini Batch (One File)	Hyperparameters To Encode the BPPR	Encode Mini Batch (All files)	Encode Mini Batch (One file)	Exe. Time server (mins)
2007-2008	[X1, X2, X3, X4, X5, X6, X7]	11094	84999	**3.2 GB**	381M	Relu, Adadelta, BS: 32, E: 10, **[2000, 1000, 500]**	**255 MB**	36M	**39.019**
		11094	84999	**3.2 GB**	381M	Relu, Adadelta, BS: 32, E: 10, **[500, 200, 100]**	**48.87 MB**	6.9M	**24.26**
	[X1, X2, X3, X4, X5, X6, X7, X8.3]	14515	84999	**4.1 GB**	499M	Relu, Adadelta, BS: 32, E: 10, **[2000, 1000, 500]**	**247.9 MB**	35M	**50.48**
		14515	84999	**4.1 GB**	499M	Relu, Adadelta, BS: 32, E: 10, **[500, 200, 100]**	**49.58 MB**	7.0M	**41.03**
Without: X6 (Rehabilitation Time)	[X1, X2, X3, X4, X5, X7]	8041	84999	**2.3 GB**	276M	Relu, Adadelta, BS: 32, E: 10, **[2000, 1000, 500]**	**255 MB**	36M	**26.43**
		8041	84999	**2.3 GB**	276M	Relu, Adadelta, BS: 32, E: 10, **[2000, 1000, 500]**	**48.87 MB**	6.9M	**17.10**
	[X1, X2, X3, X4, X5, X7, X8.3]	11462	84999	**3.3 GB**	394M	Relu, Adadelta, BS: 32, E: 10, **[2000, 1000, 500]**	**255 MB**	36M	**37.86**
		11462	84999	**3.3 GB**	394M	Relu, Adadelta, BS: 32, E: 10, **[2000, 1000, 500]**	**50.27 MB**	7.1M	**29.66**

		Mini Batch (Files)
Data Training	**84999**	**9**
Data Training	72000	6
Data Training-Dev	24999	3

Stage 3. dIAgnoseNet Supervised Learning

Training the BPPR and LPPR by Random Forest Algorithm

To Classify the Inpatients Major Clinical Category

At Intensive Care Unit in PACA Region.

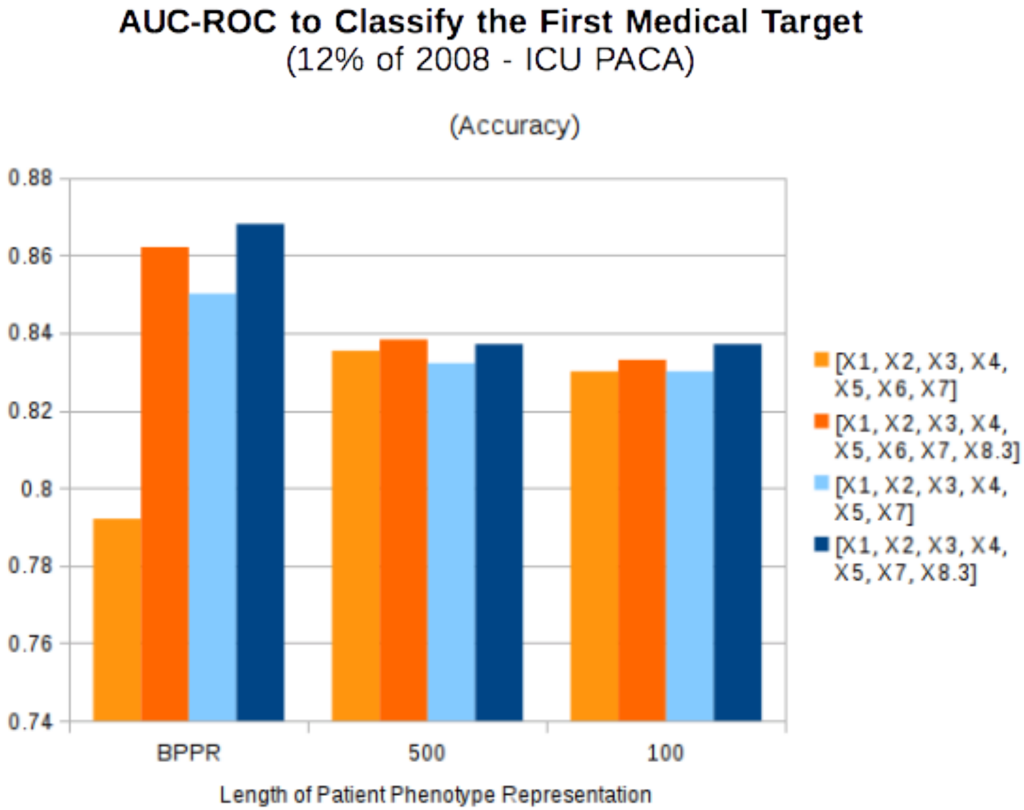
EHR Common Version	Features Composition	BPPR (Features)	Patients (Samples)	Features Mini Batch (All files)	No. Labels (Y1)	Labels (Y1) Mini Batch (All files)	Random Forest (AUC–ROC)	Exe. Time server (mins)
2007-2008	[X1, X2, X3, X4, X5, X6, X7]	11094	84999	3.2 GB	23	21 MB	0.792	10.53
		500	84999	255 MB	23	21 MB	0.8353	3.55
		100	84999	48.87 MB	23	21 MB	0.83	3.312
	[X1, X2, X3, X4, X5, X6, X7, X8.3]	14515	84999	4.1 GB	23	21 MB	0.862	10.29
		500	84999	247.9 MB	23	21 MB	0.8382	3.72
		100	84999	49.58 MB	23	21 MB	0.833	3.74
2007-2008 without reeduca-tion time	[X1, X2, X3, X4, X5, X7]	8041	84999	2.3 GB	23	21 MB	0.85	7.0
		500	84999	255 MB	23	21 MB	0.8321	3.66
		100	84999	48.87 MB	23	21 MB	0.83	3.97
	[X1, X2, X3, X4, X5, X7, X8.3]	11462	84999	3.3 GB	23	21 MB	0.868	8.80
		500	84999	255 MB	23	21 MB	0.837	3.73
		100	84999	50.27 MB	23	21 MB	0.837	3.44

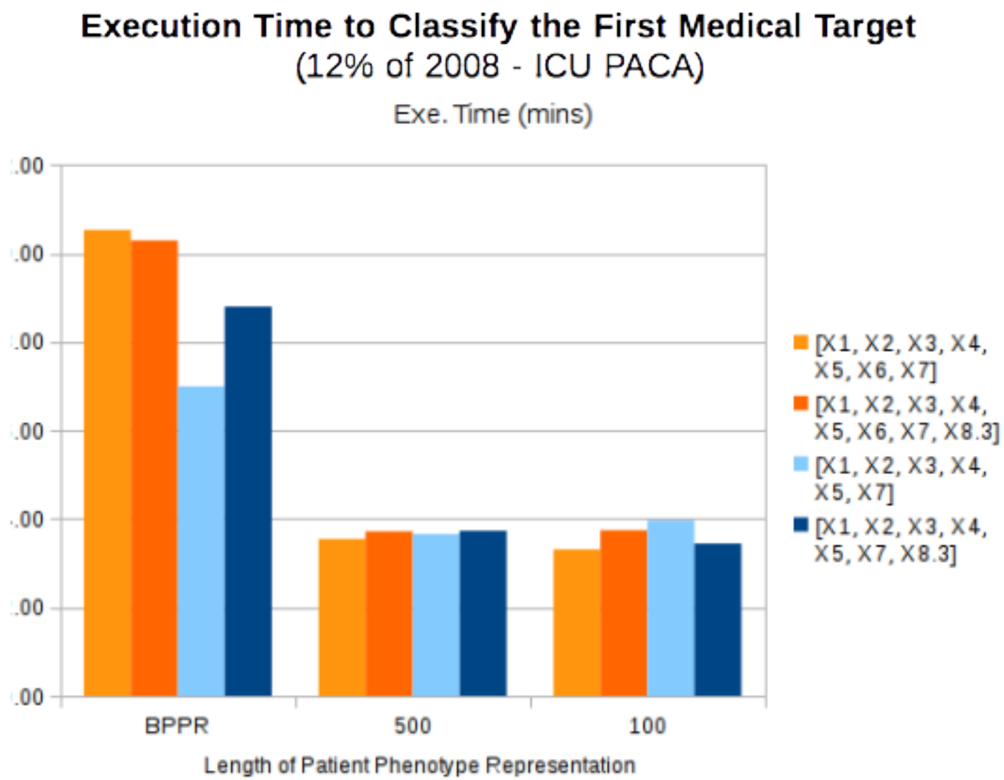
Summary Results to Predict the *'Major Clinical Category'*

1)Reeducation time do not help us for classification

2) Auto-encoding stage do not modify the accuracy

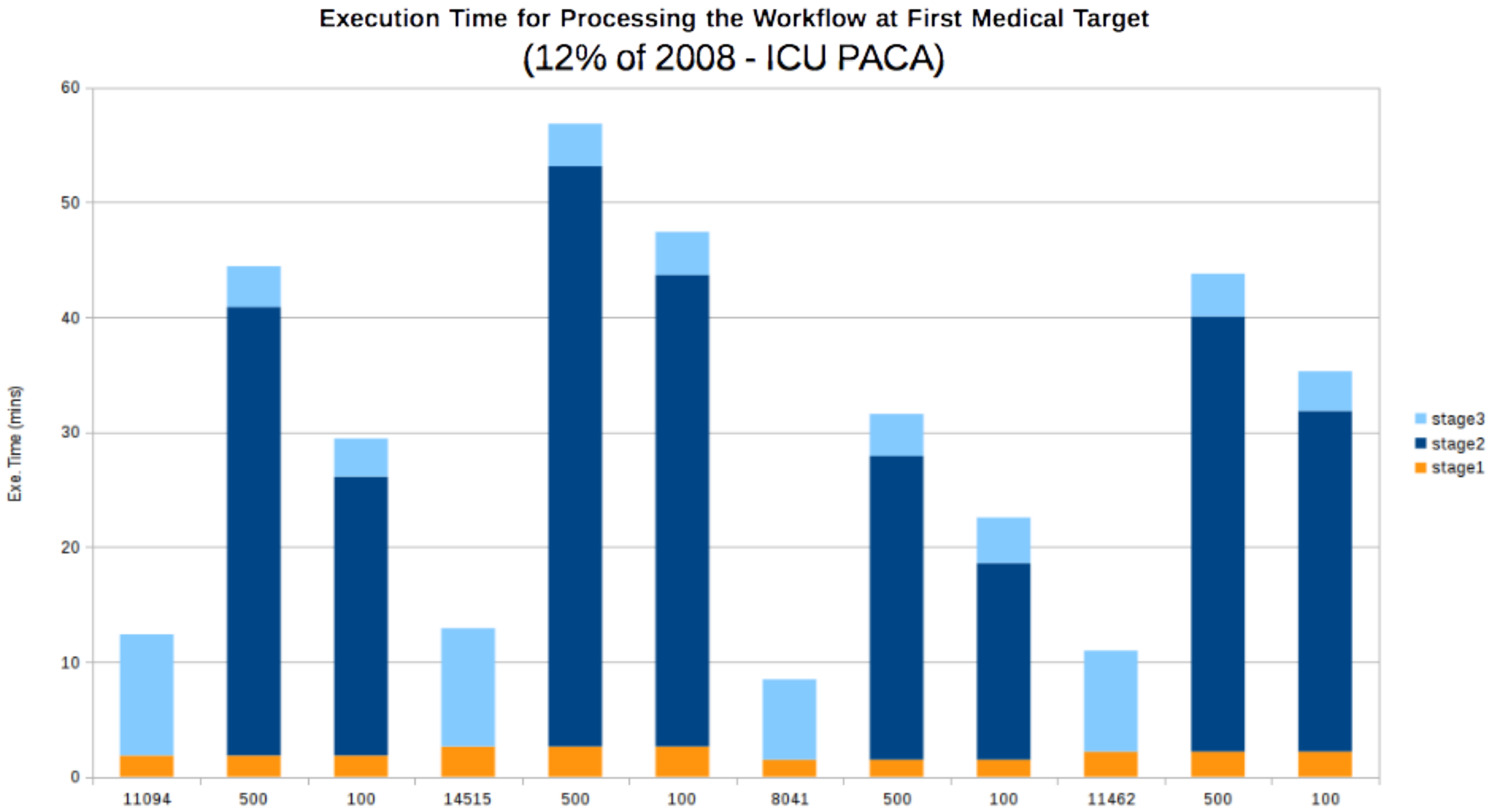
3) Auto-encoding stage reduce the classification time



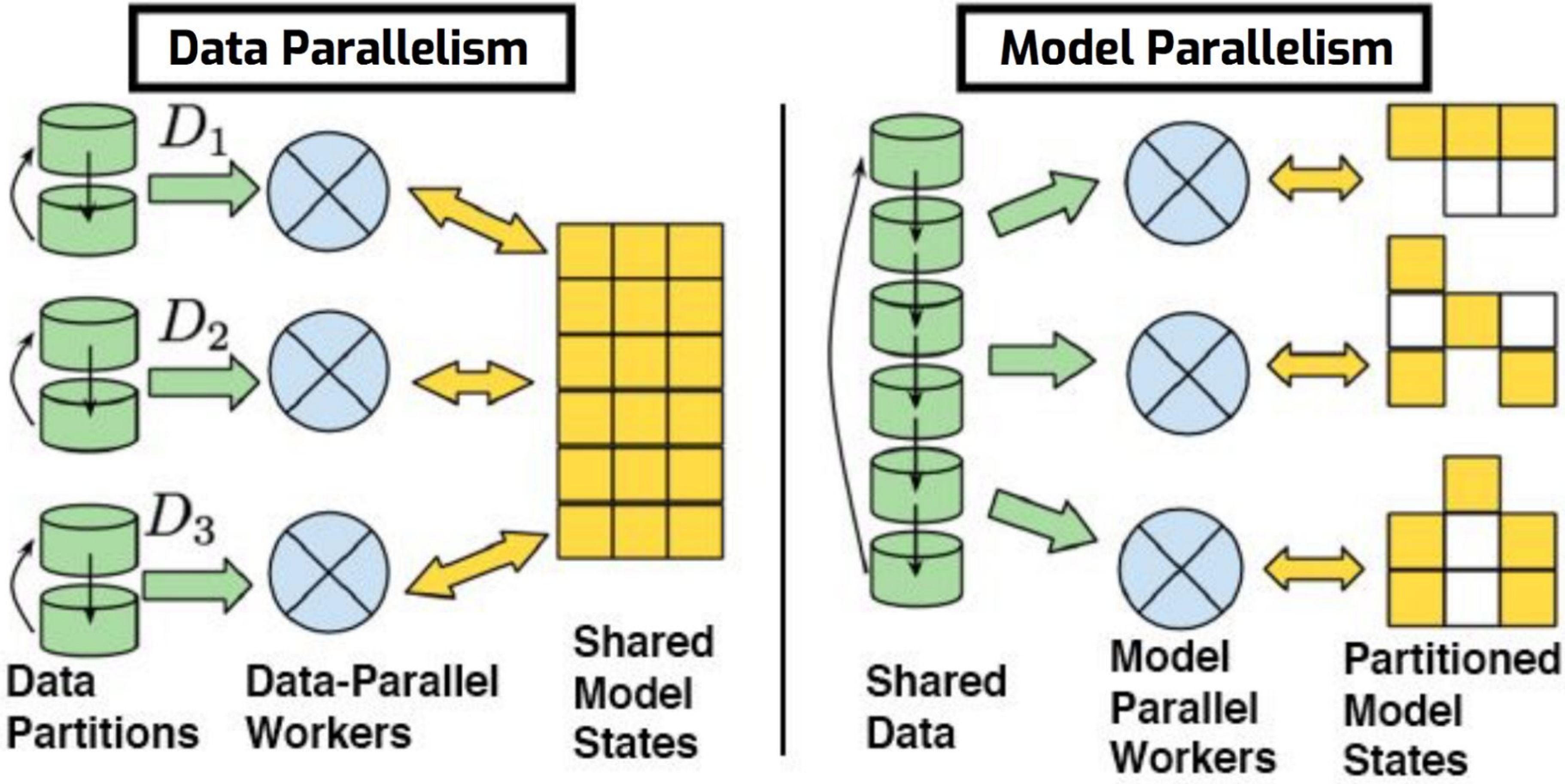


Summary Results to Predict the *'Major Clinical Category'*

3) But auto-encoding stage is very long



How can we train large, powerful models fastly?



1. Adapted from Large Scale Distributed Deep Networks. By Jeffrey Dean et al. NIPS 2012

How can we train large, powerful models fastly?

Exploit many kinds of parallelism

1. Data parallelism

Fit large amout of data

+ Speeds up the training.

+ Good for forward pass (independent)

- I/O Intensive.

  - Backpropagation requires all-to-all communication only when accumulating results

- Requires allocation of all parameters on each processor

+ Synchronous

  N replicas equivalent to an N times larger batch size

  Pro: No noise

  Con: Less fault tolerant (requires some recovery if any single machine fails)

vs. Asynchronous.

  Con: Noise in gradients

  Pro: Relatively fault tolerant (failure in model replica doesn't block other replicas)

1. Adapted from Large Scale Distributed Deep Networks. By Jeffrey Dean et al. NIPS 2012

How can we train large, powerful models fastly?

Exploit many kinds of parallelism

2. **Model parallelism**

   Fit large model

   \+ Parameters can be divided across processors

   \- Mini-batch has to be copied to all processors

   \- Back propagation requires all-to-all communication every layer

   \+ Distributed execution of tasks

   \+ Task scheduling on computational resources

3. Or **Hybrid approach** : M asynchronous groups of N synchronous replicas

On top of distributed architecture

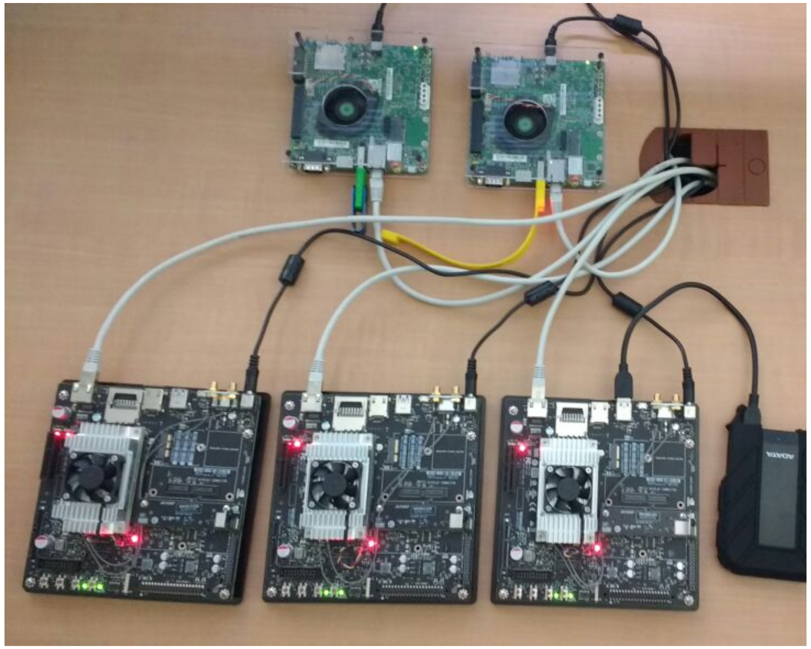
- Cluster of Jetson

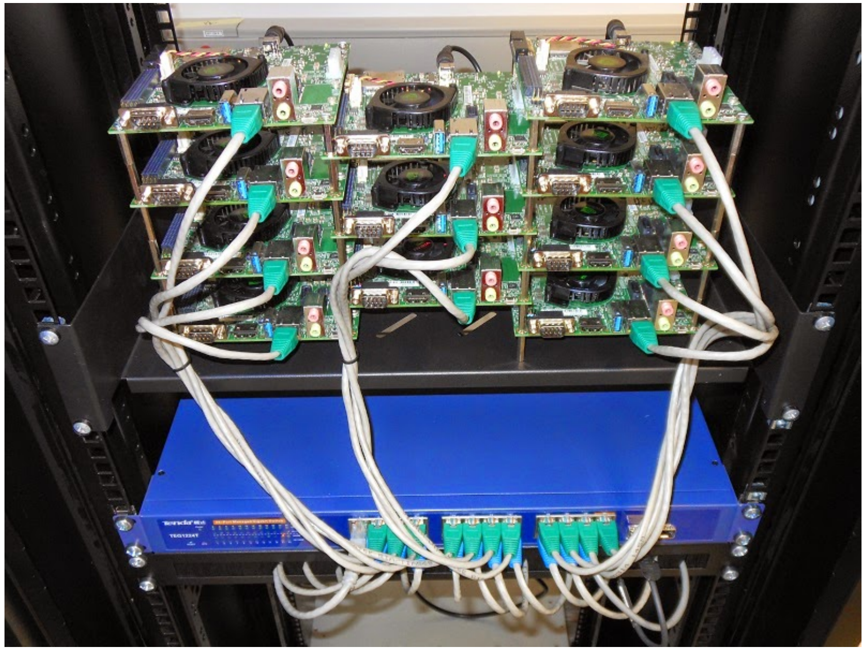
In order to produce a predictive model

→ Which approach will be better for a given dataset size + neural network

1. Adapted from Large Scale Distributed Deep Networks**.** By Jeffrey Dean et al. NIPS 2012

Jetson Cluster





Towards a Green Intelligence Applied to Health Care and Well-Being

John A. GARCÍA. H.

Frédéric PRECIOSO, Pascal STACCINI, Michel RIVEILL

Laboratoire d'Informatique, Signaux et Systèmes de Sophia Antipolis - I3S