Challenges of Learning for Healthcare

Assaf Schuster, Computer Science, Technion

Motivation

nature
International journal of science

Letter | Published: 25 January 2017

Dermatologist-level classification of skin cancer with deep neural networks

Andre Esteva, Brett Kuprel, Roberto A. Novoa, Justin Ko, Susan M. Swetter, Helen M. Blau & Sebastian Thrun

Convolutional NN for images

LARGE dataset

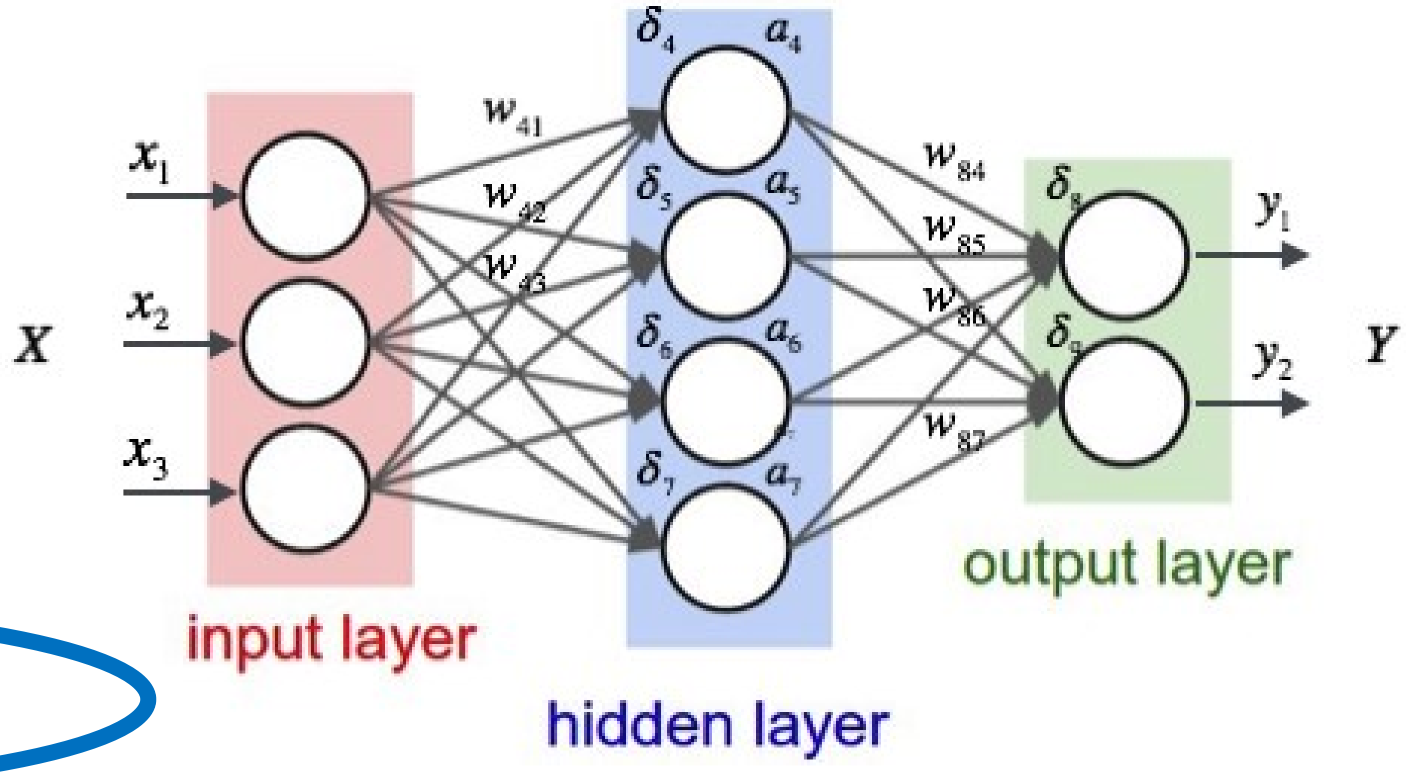
Labels==experts time

many fine-grained object categories[6,7,8,9,10,11]. Here we demonstrate classification of skin lesions using a single CNN, trained end-to-end from images directly, using only pixels and disease labels as inputs. We train a CNN using a dataset of 129,450 clinical images—two orders of magnitude larger than previous datasets[12]—consisting of 2,032 different diseases. We test its performance against 21 board-certified

The Neural Network Revolution

Advantages:
- Accuracy
- Featureless

Issues:
- Require BIG DATA
- Featureless
- Architecture zoo
- Hyperparameters desert
- Computational complexity



Hyperparameters

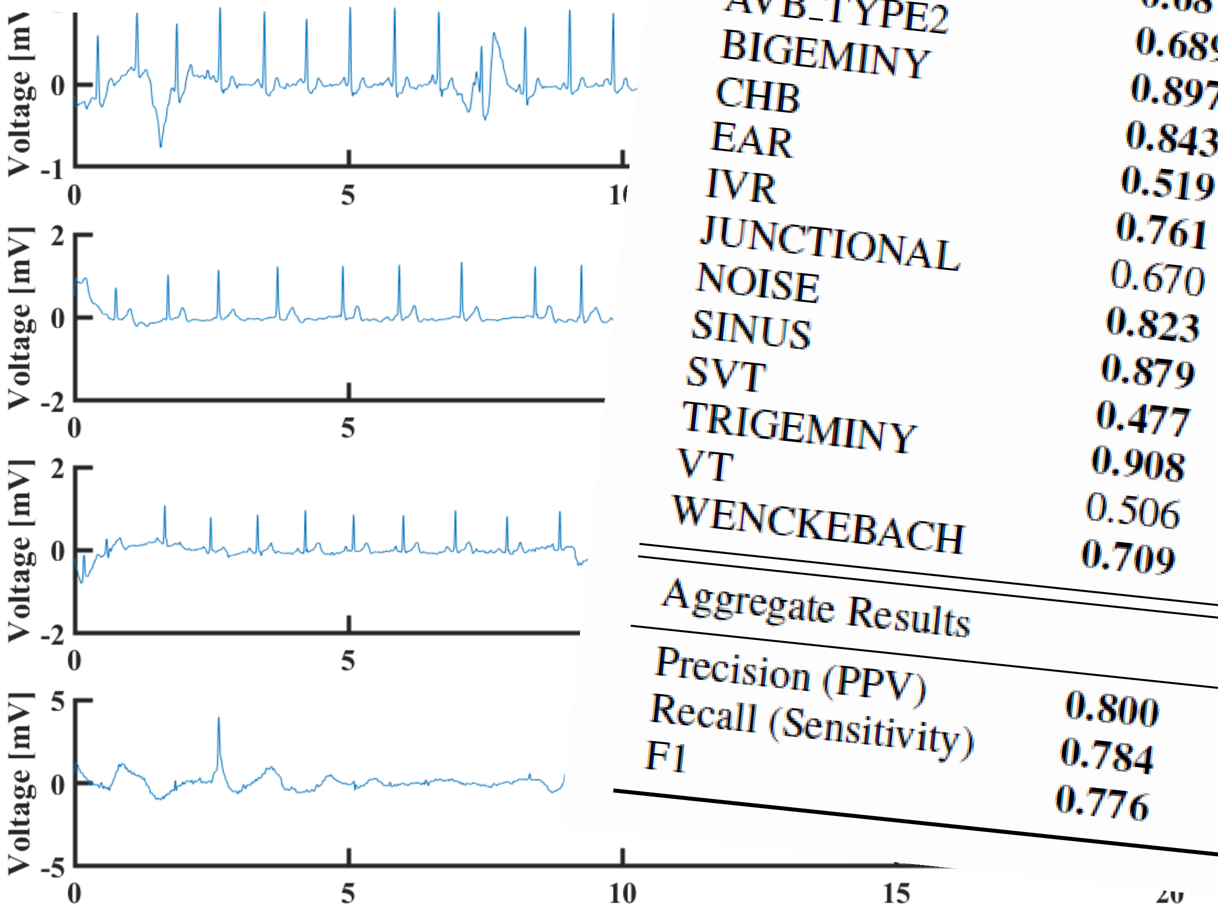
Dermatologist-level classification of skin cancer with deep neural networks

Andre Esteva[1*], Brett Kuprel[1*], Roberto A. Novoa[2,3], Justin Ko[2], Susan M. Swetter[2,4], Helen M. Blau[5] & Sebastian Thrun[6]

Our CNN is trained using backpropagation. All layers of the network are fine-tuned using the same global learning rate of 0.001 and a decay factor of 16 every 30 epochs. We use RMSProp with a decay of 0.9, momentum of 0.9 and epsilon of 0.1. We use Google's TensorFlow[30] deep learning framework to train, validate and test our network. During training, images are augmented by a factor of 720. Each image is rotated randomly between 0° and 359°. The largest upright inscribed rectangle is then cropped from the image, and is flipped vertically with a probability of 0.5.

Automatic ... ses

Gliner & Ya ...)



Class-level F1 Score	Seq Model	Seq Cardiol.	Set Model	Set Cardiol.
AFIB	0.604	0.515	0.667	0.544
AFL	0.687	0.635	0.679	0.646
AVB_TYPE2	0.689	0.535	0.656	0.529
BIGEMINY	0.897	0.837	0.870	0.849
CHB	0.843	0.701	0.852	0.685
EAR	0.519	0.476	0.571	0.529
IVR	0.761	0.632	0.774	0.720
JUNCTIONAL	0.670	0.684	0.783	0.674
NOISE	0.823	0.768	0.704	0.689
SINUS	0.879	0.847	0.939	0.907
SVT	0.477	0.449	0.658	0.556
TRIGEMINY	0.908	0.843	0.870	0.816
VT	0.506	0.566	0.694	0.769
WENCKEBACH	0.709	0.593	0.806	0.736
Aggregate Results				
Precision (PPV)	0.800	0.723	0.809	0.763
Recall (Sensitivity)	0.784	0.724	0.827	0.744
F1	0.776	0.719	0.809	0.751

Normal

Atrial fibrillation

Other rhythm

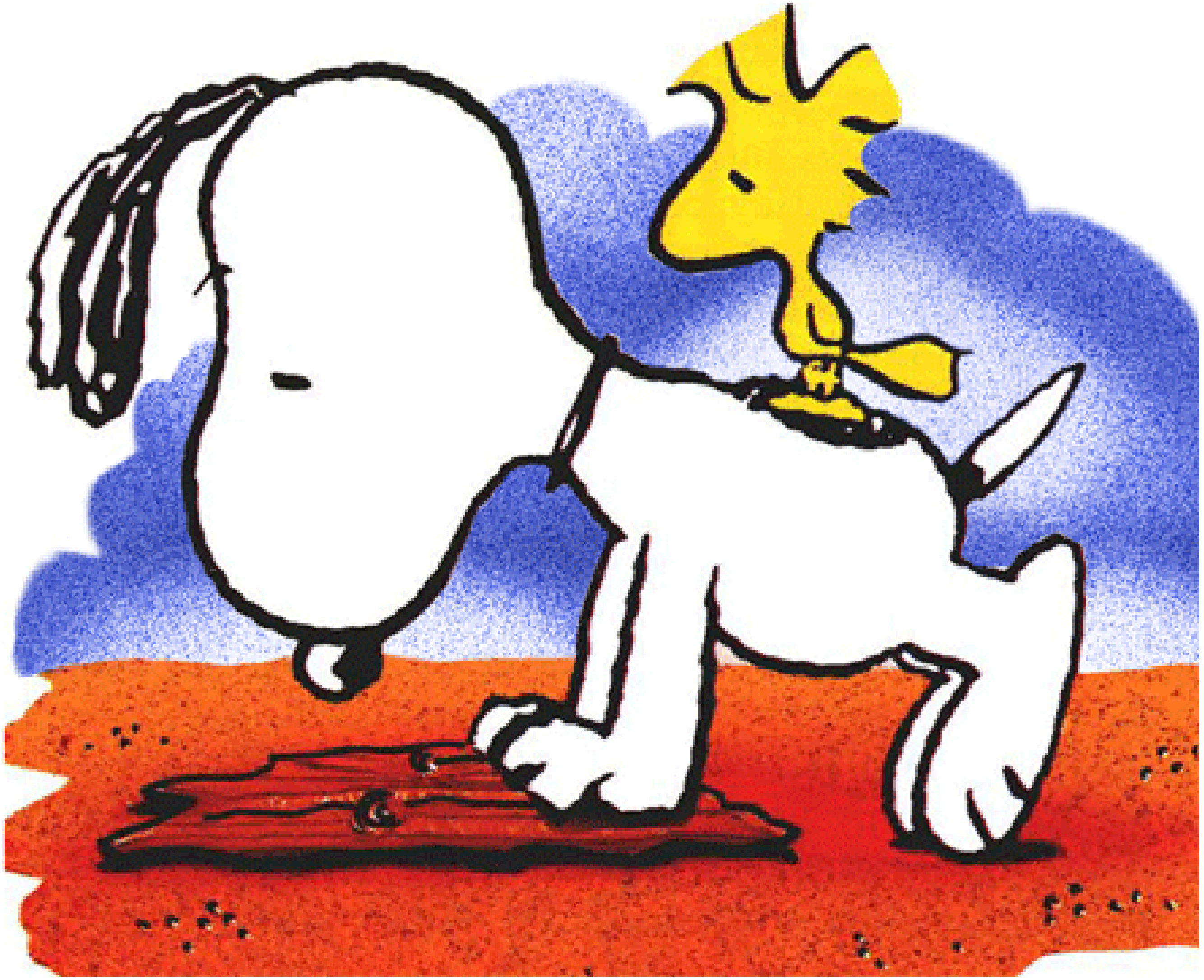
Noise

Andrew Ng group, Nature medicine 2018 (tens of thousands examples)

Challenges
in applying modern ML
to healthcare

Challenge 0: The Quest for Data

- For training.....



Data acquisition issues

Privacy:

- General Data Protection Regulation – GDPR, HIPPA
- What is data anonymization?
- Managers: better keep the data locked in the safe…

Labelled data:

- Experts time is valuable
- Nature paper used 21 dermatologists (on a subset of dataset)

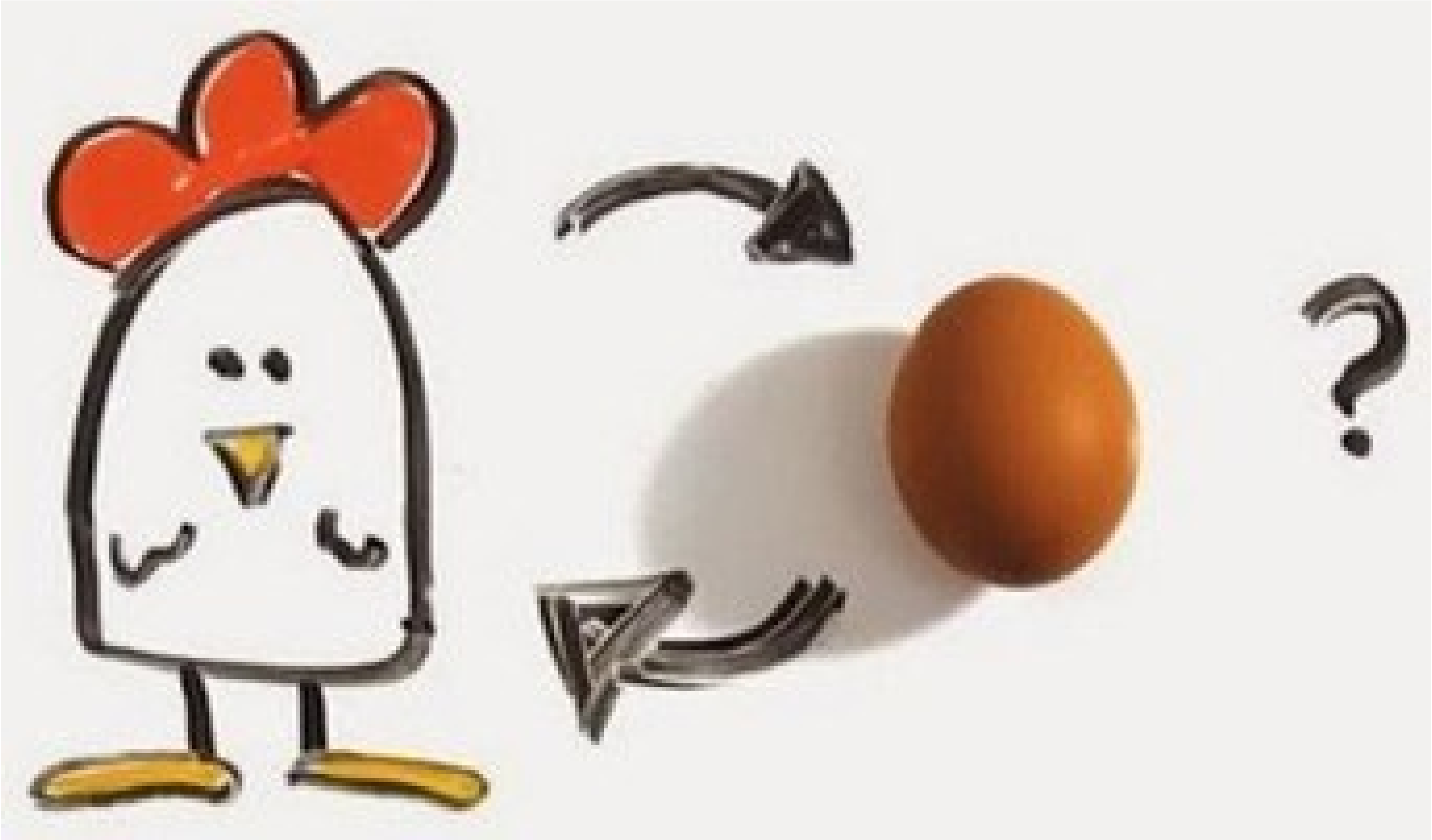
Biomedical Engineers Awareness:

- Current defibrillators collect and transmit only a few seconds before/after a VF event
- Suppose you'd like to… predict the cardiac event minutes before it happen

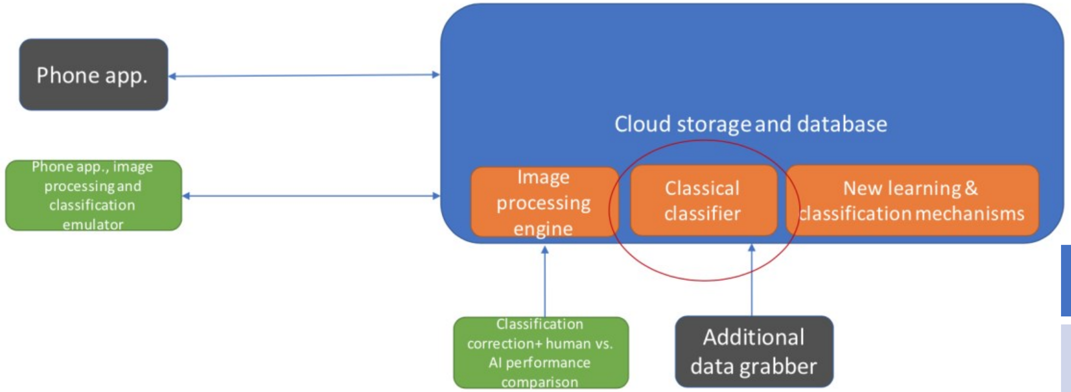
Recording patients:

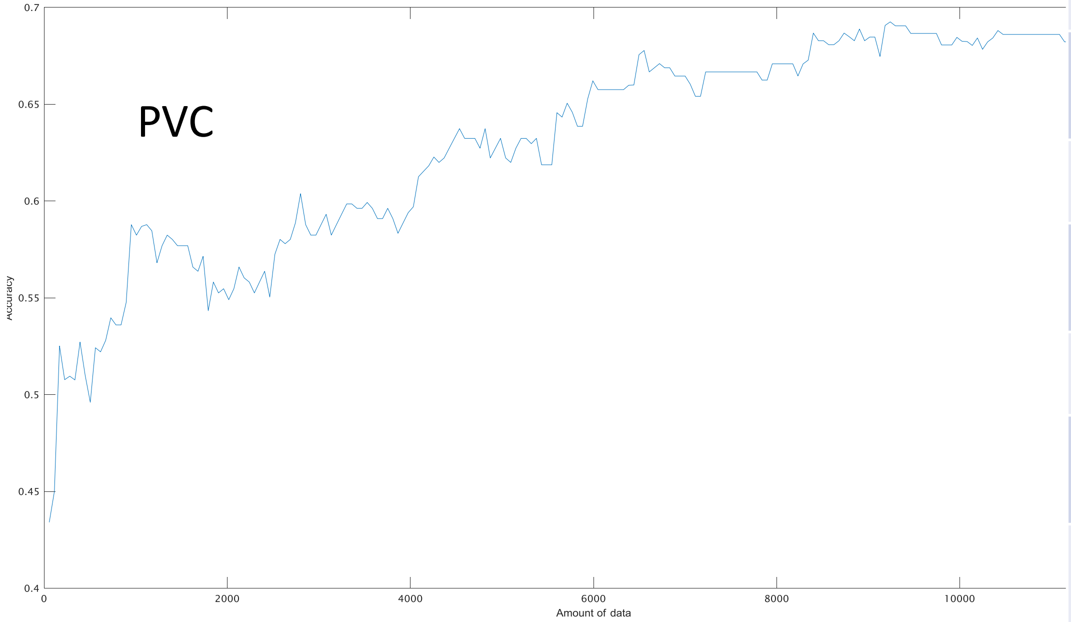
- Scarce resource
- Non homogenic
- Complicated rec. at home
- Rare visits to lab
- Expensive devices
- Etc.

Data Acquisition - the egg and the chicken

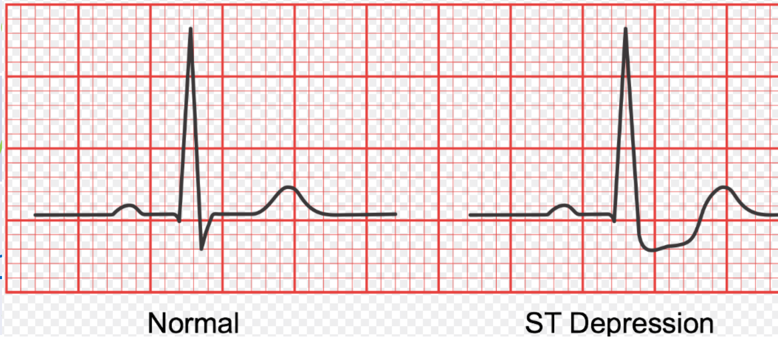


Automatic disease classification

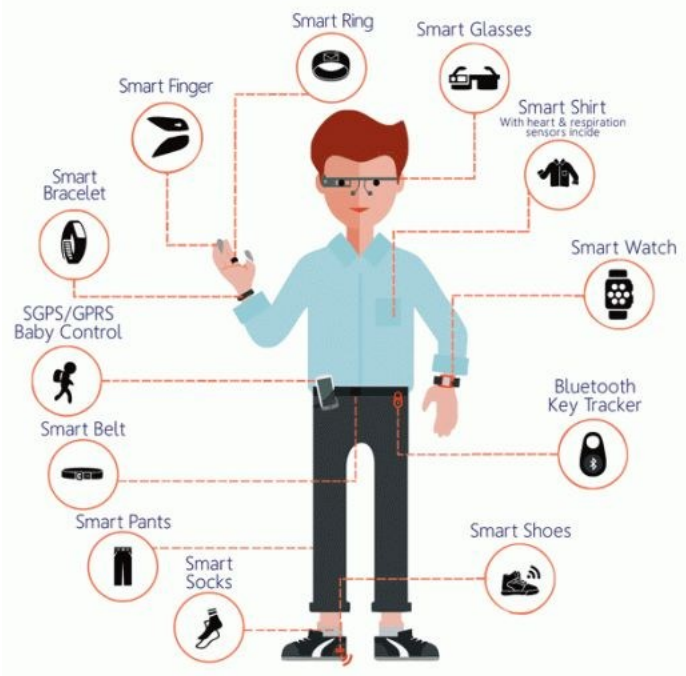




Disease	Dominant symptoms
AF	HRV
I-AVB	PR Interval>300ms
LBBB	QRS duration>120ms, No Q wave on lateral lead (V6), M shaped R wave at V6, W shaped R wave at V1
RBBB	T wave inversion in V1, ‘M’ QRS complex at V1, possible W shape at V6
PAC	PR interval shorter, QRS normal, P wave abnormally shaped, two types of R-R intervals
PVC	QRS wider, QRS bizarre, no prec
ST-D	Vertical distance between the p 1 mm in V5-V6, or 1.5 mm in AV
ST-E	The vertical distance inside the after the J-point is at least 0.1 m a limb lead or 0.2 mV (2 mm or



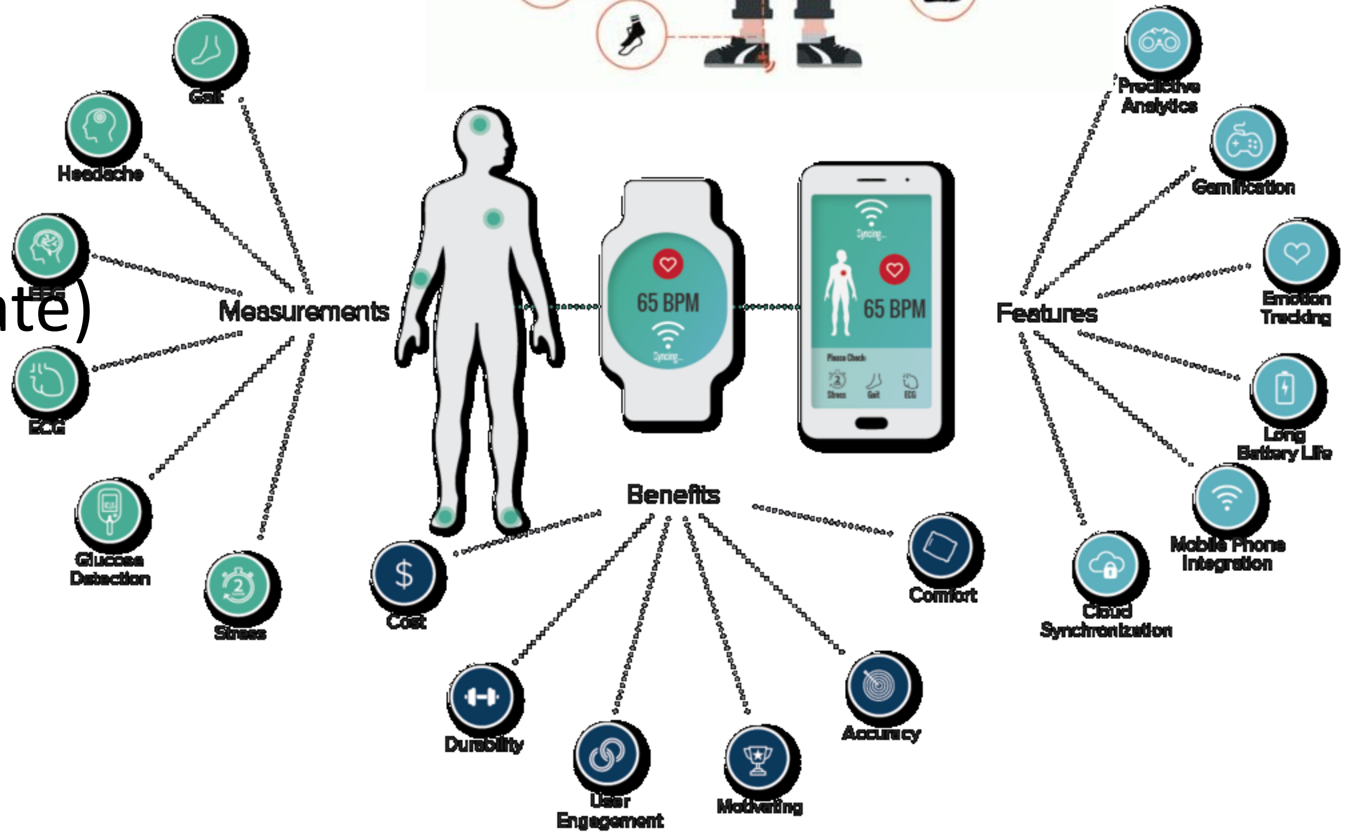
The IoT Revolution I – Monitoring Devices



- Source of BIG DATA
- Source of Rich Data

Issues:

- Battery (compute+communicate)
- Connectivity
- Resources
- Privacy

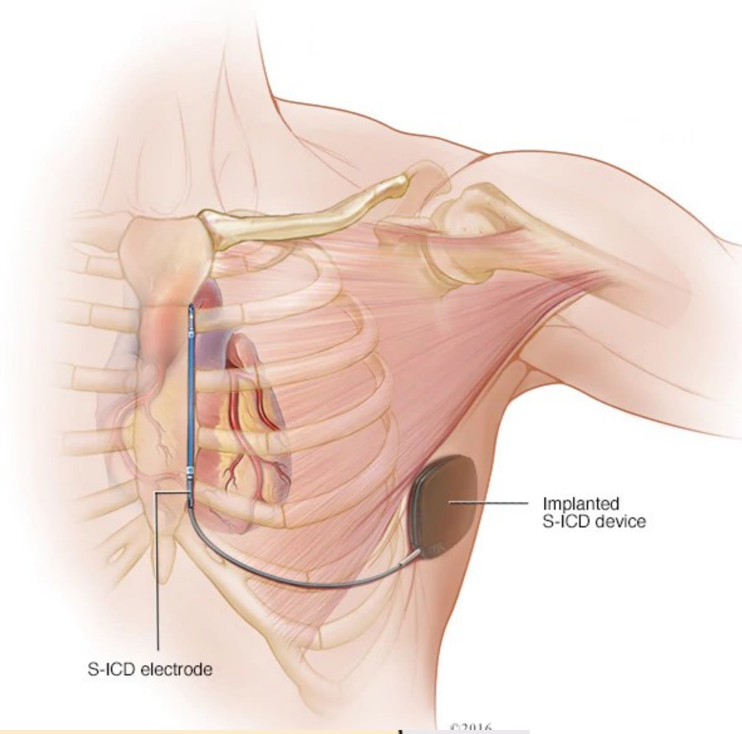


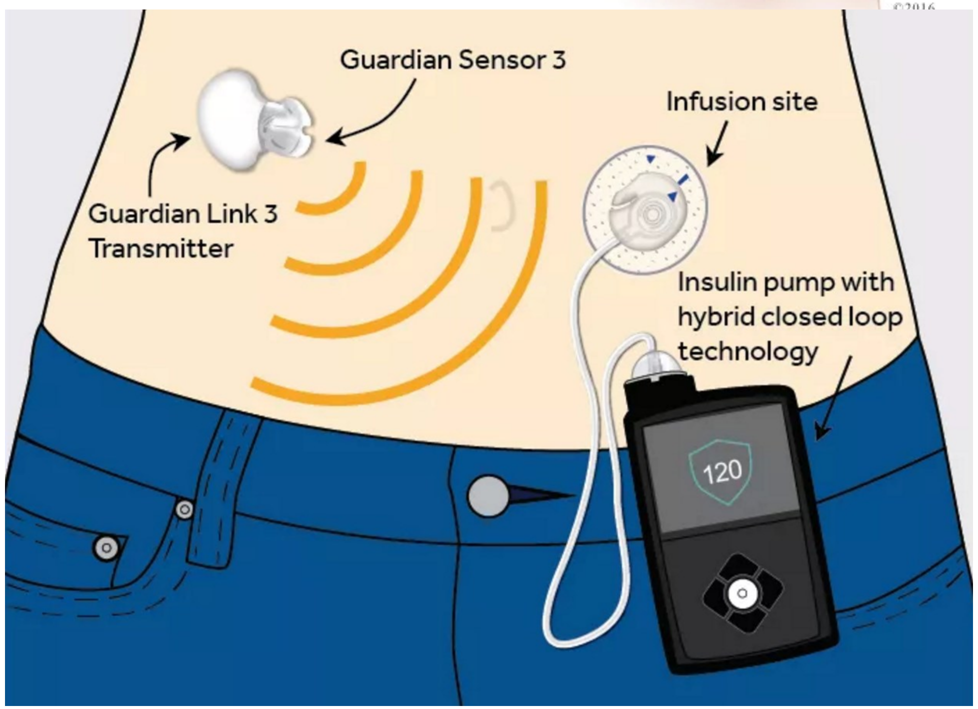
The IoT Revolution II - Life-Saving Devices

Additional issues:

- Autonomy
- Edge computing
- Collaboration with the cloud

- Security



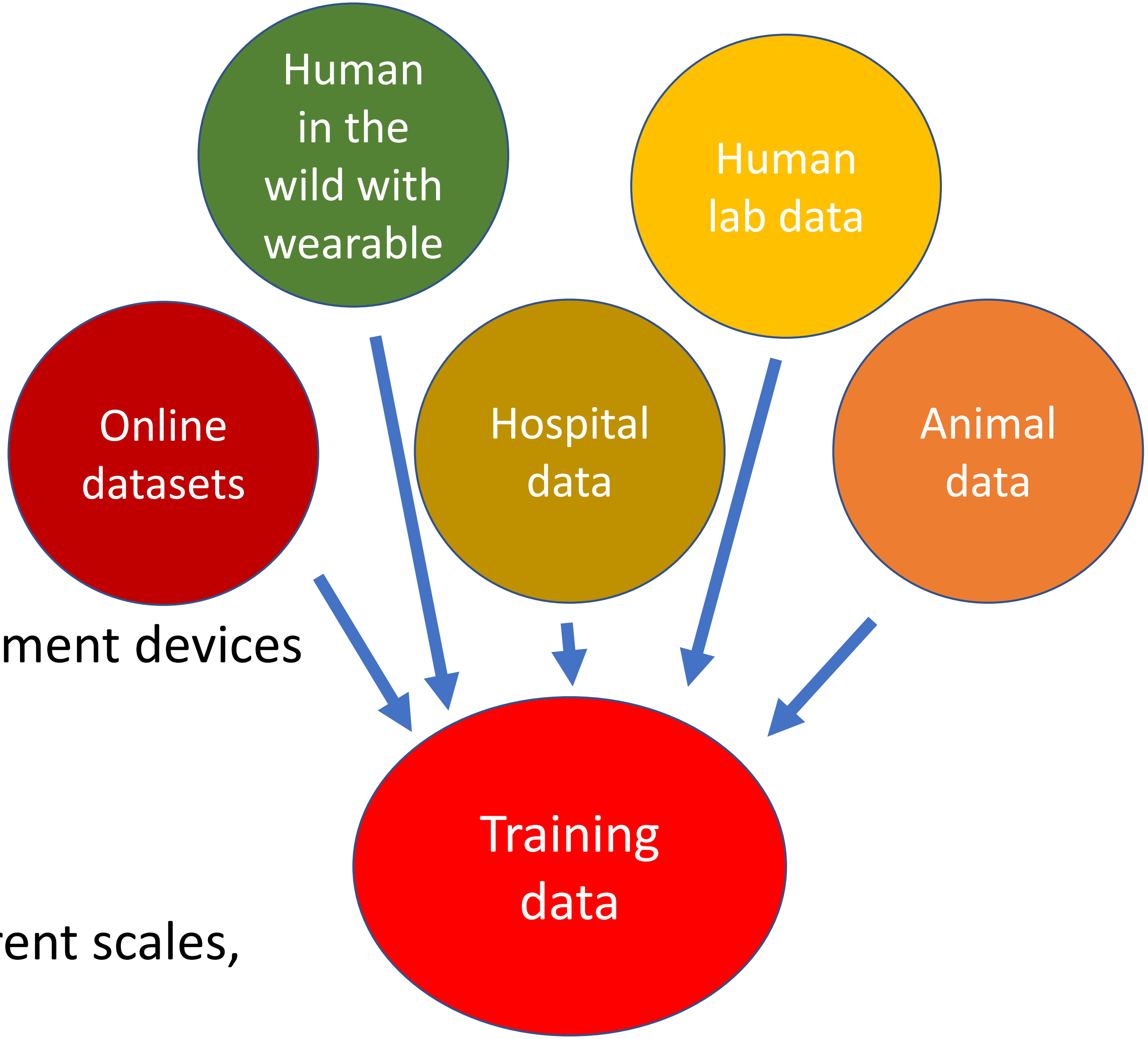


Challenge I: Heterogeneous Data Sources

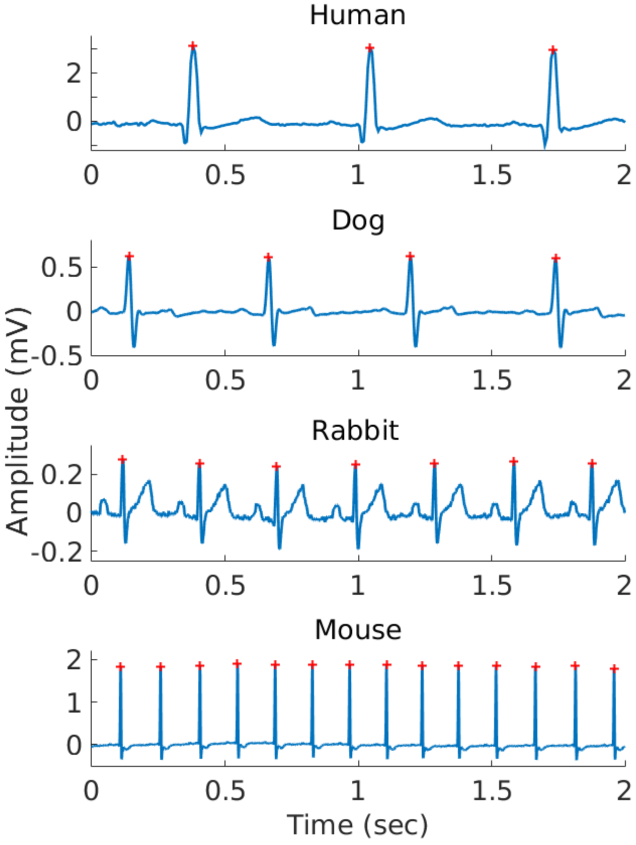
Data collected:

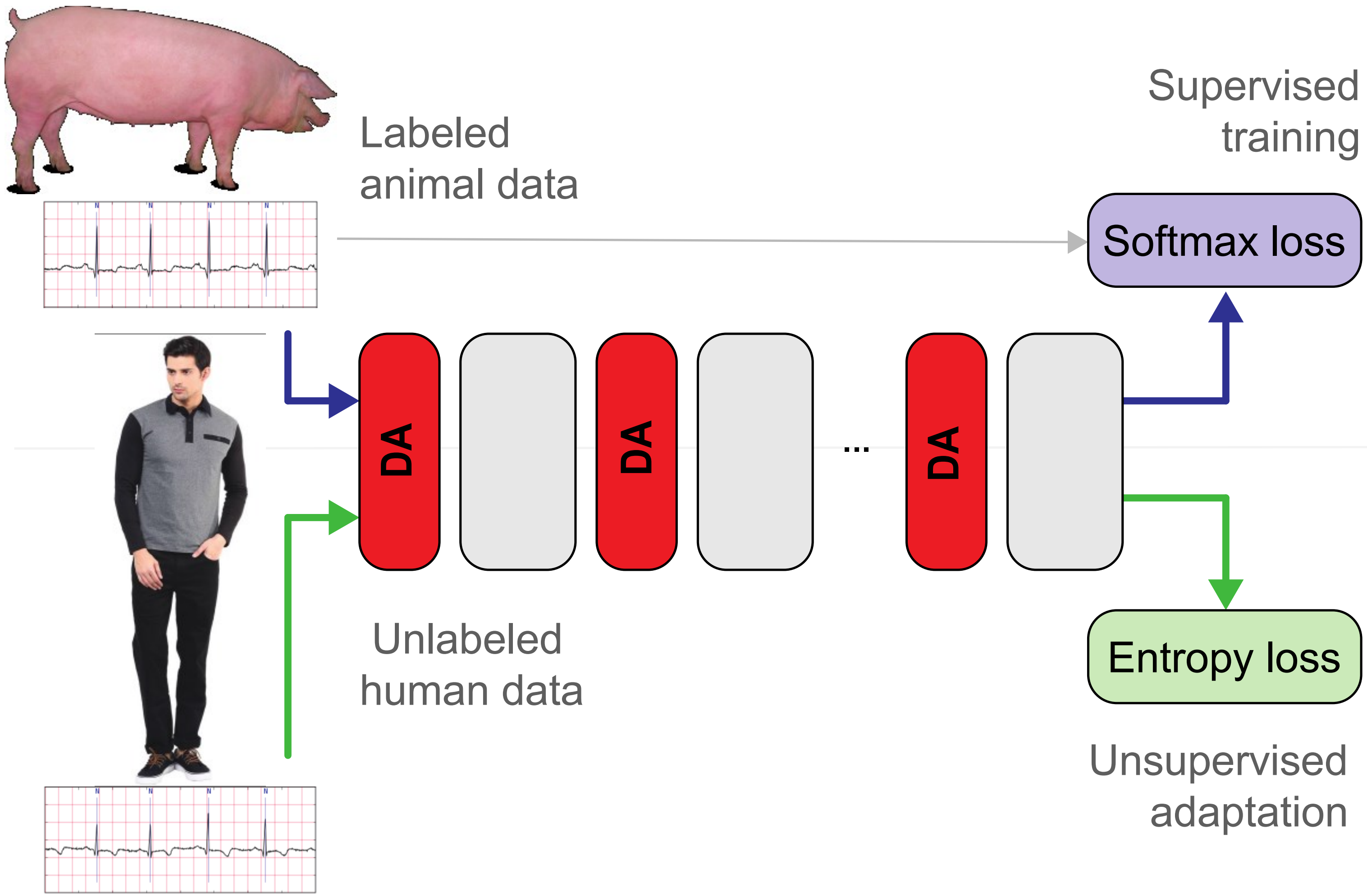
- On different animals
- Using different measurement devices
- At different times
- By different people

➔ Different ranges, different scales, shifts, drifts, …..



Domain Adaptation/Transfer





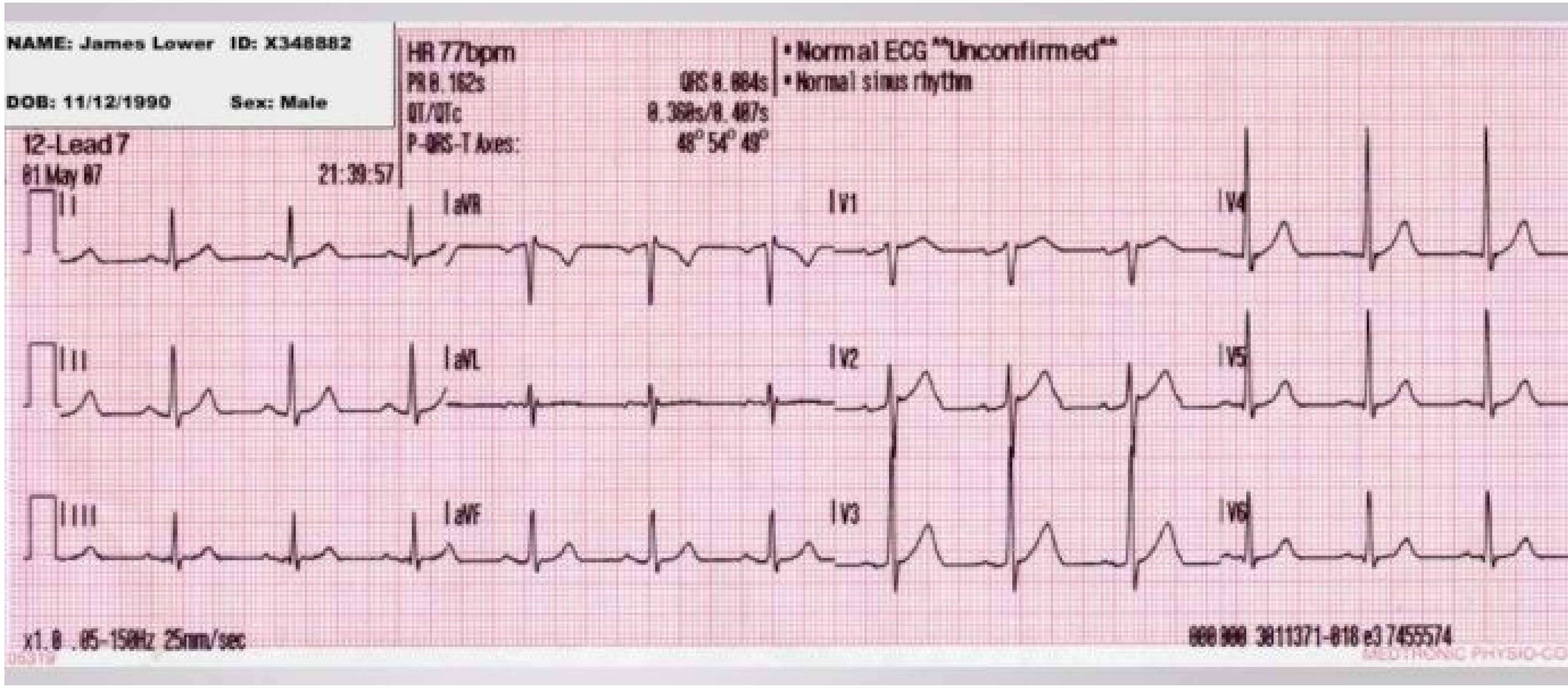
Transfer learning

Dermatologist-level classification of skin cancer with deep neural networks

Andre Esteva[1*], Brett Kuprel[1*], Roberto A. Novoa[2,3], Justin Ko[2], Susan M. Swetter[2,4], Helen M. Blau[5] & Sebastian Thrun[6]

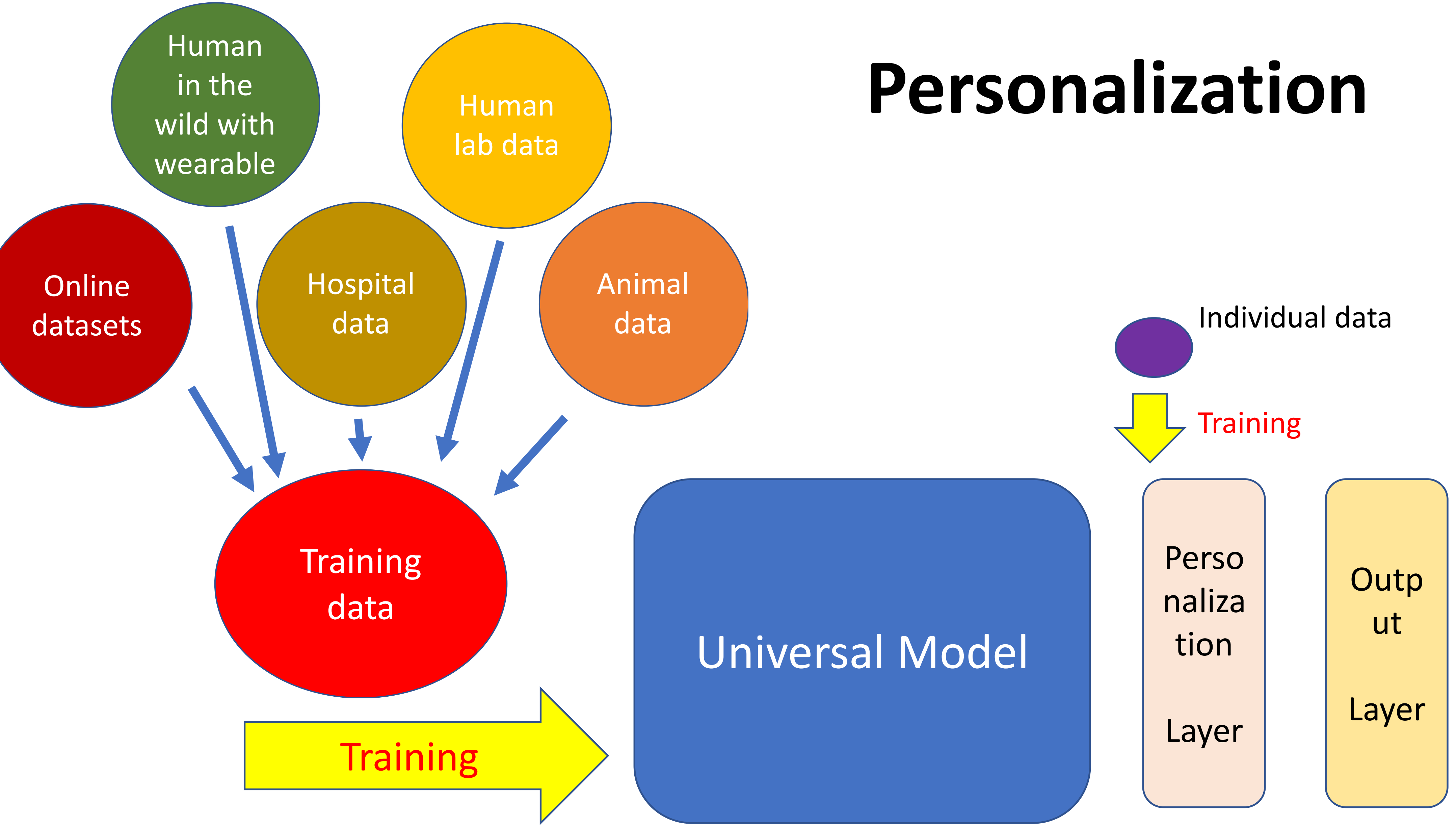
Training algorithm. We use Google's Inception v3 CNN architecture pretrained to 93.33% top-five accuracy on the 1,000 object classes (1.28 million images) of the 2014 ImageNet Challenge following ref. 9. We then remove the final classification layer from the network and retrain it with our dataset, fine-tuning the parameters across all layers. During training we resize each image to 299 × 299 pixels in order to make it compatible with the original dimensions of the Inception v3 network architecture and leverage the natural-image features learned by the ImageNet

Challenge II: Small Personal Data



- Example: prediction of Atrial/Ventrical Fibrillation
- Must compare with individual-specific characteristics
- Personal data is …. Well …. tiny

Personalization



Afimilk data - Cowlization

- ~30,000 dairy barns, x00-x000s cows in each

- 5 sensors on every cow
- Milk quality (e.g., fat), every 200cc
- Weather
- Food
- Genetic information
- Etc.

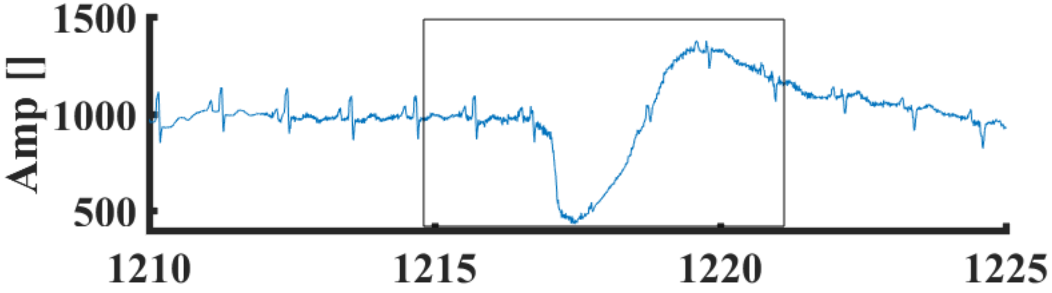
- Predict cow health
- Milk quality

Challenge III: Noise

Mobile ECG example

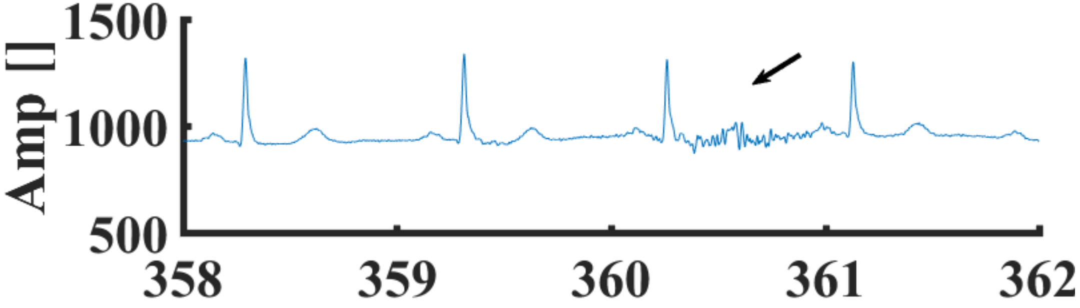
Noise modeling

Sudden movement



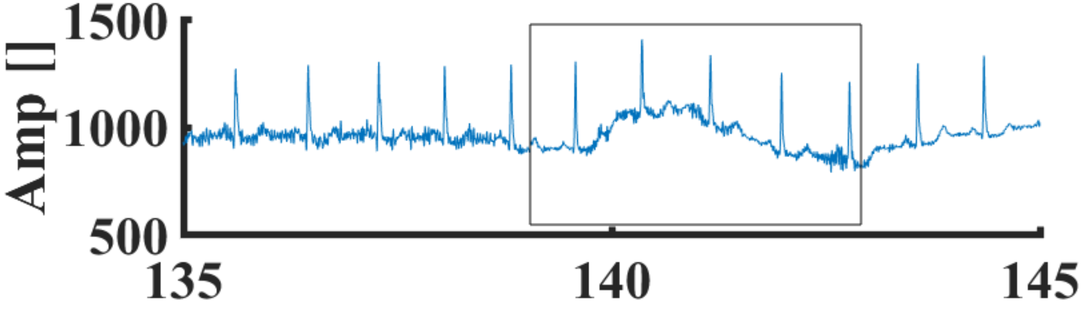
Smoothing, signal processing

Environmental noise



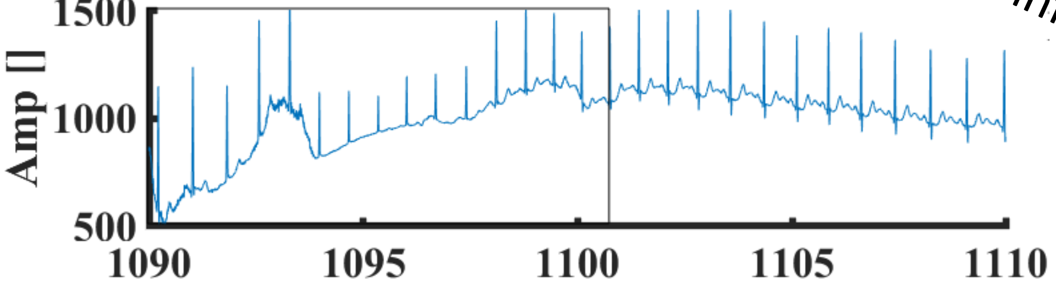
Domain knowledge

Breathing oscillations



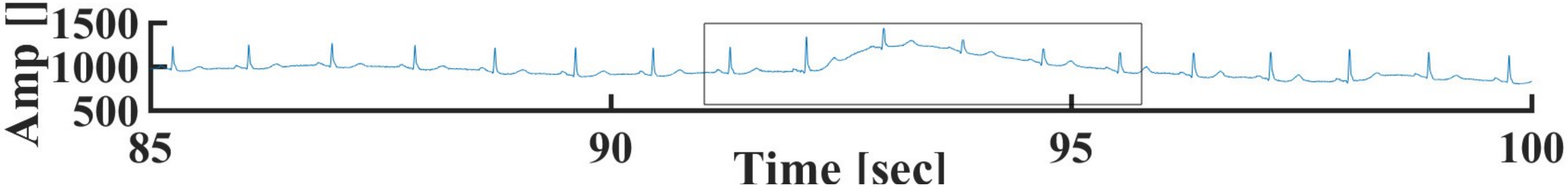
Correlating external information

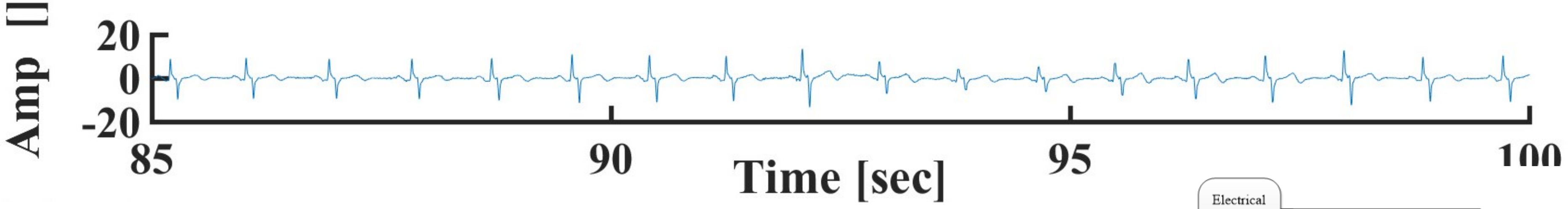
Electrical

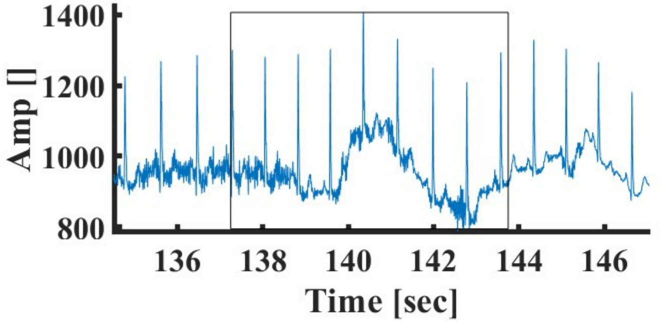


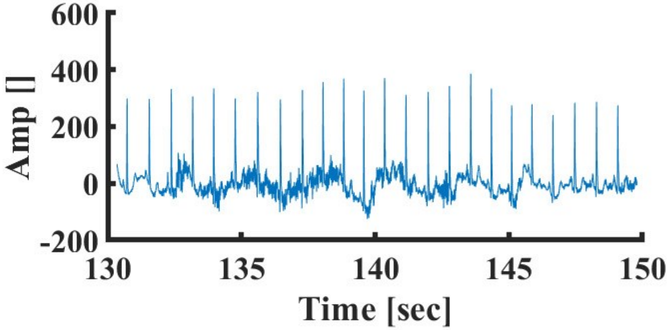
Solution for noise problem

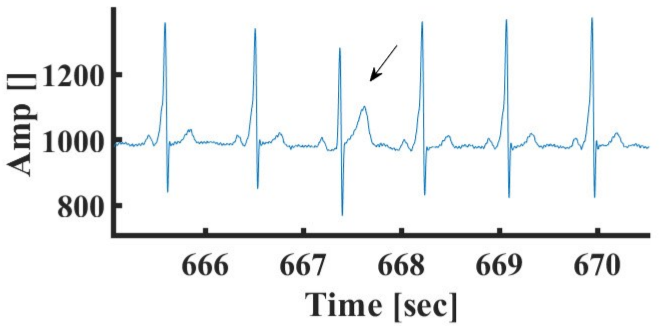
Mobile ECG example

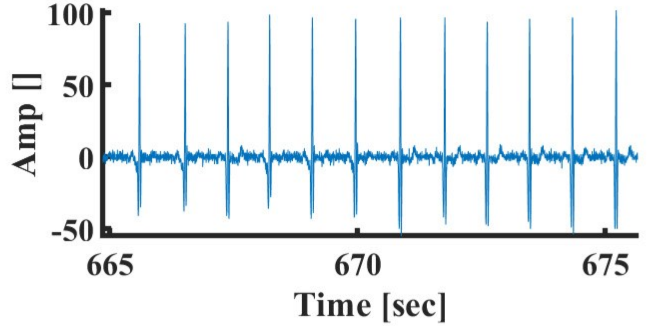


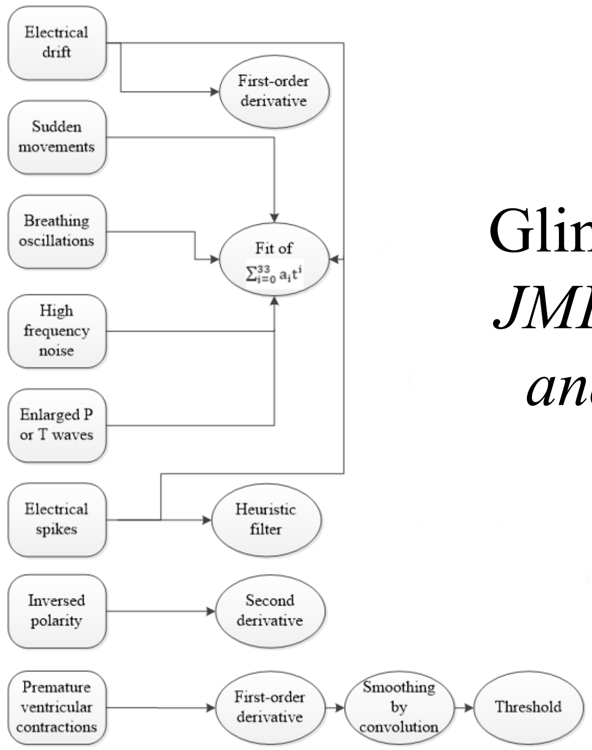






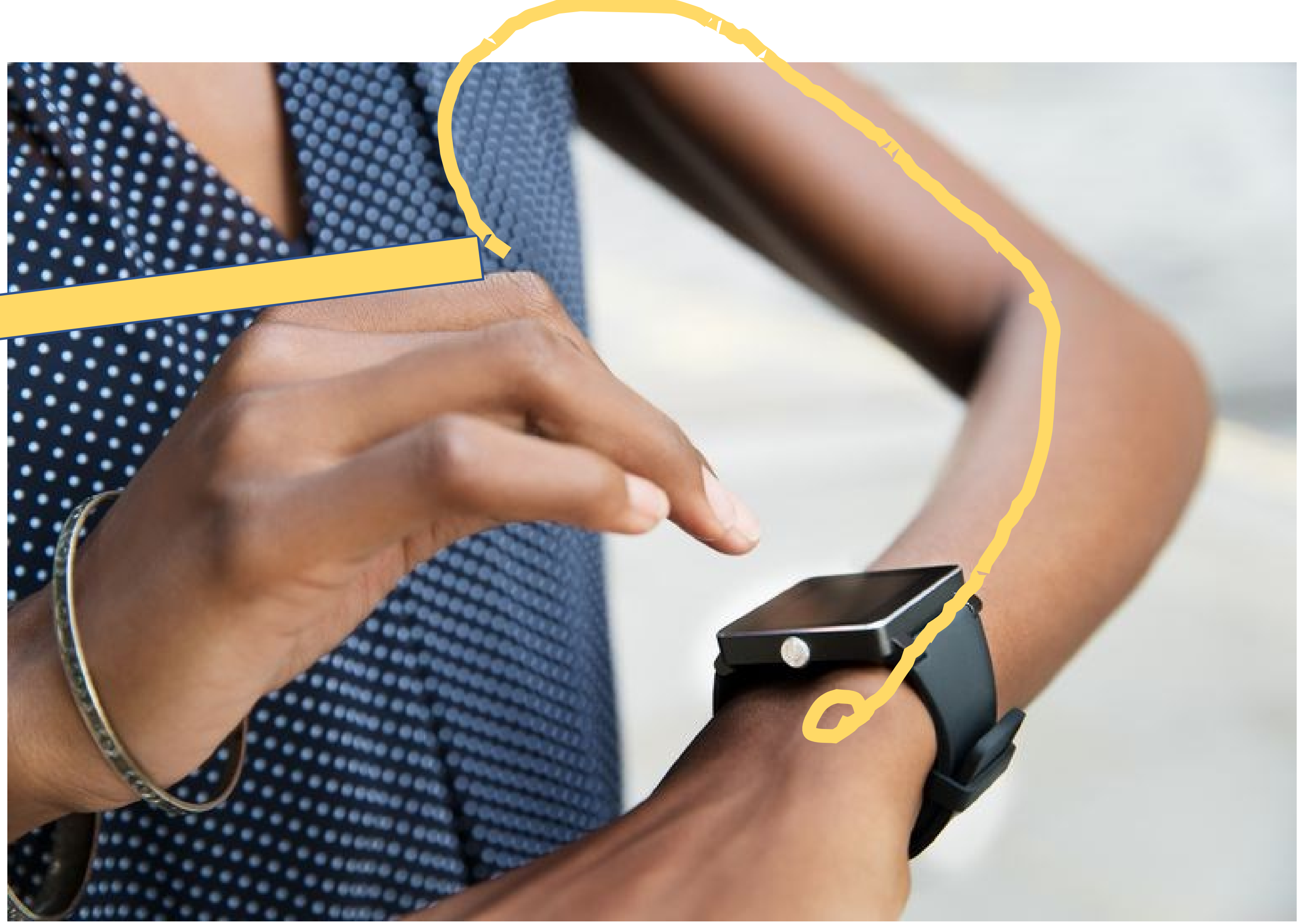






Gliner & Yaniv
JMIR mHealth and uHealth
2018

Other sources of noise – PPG example

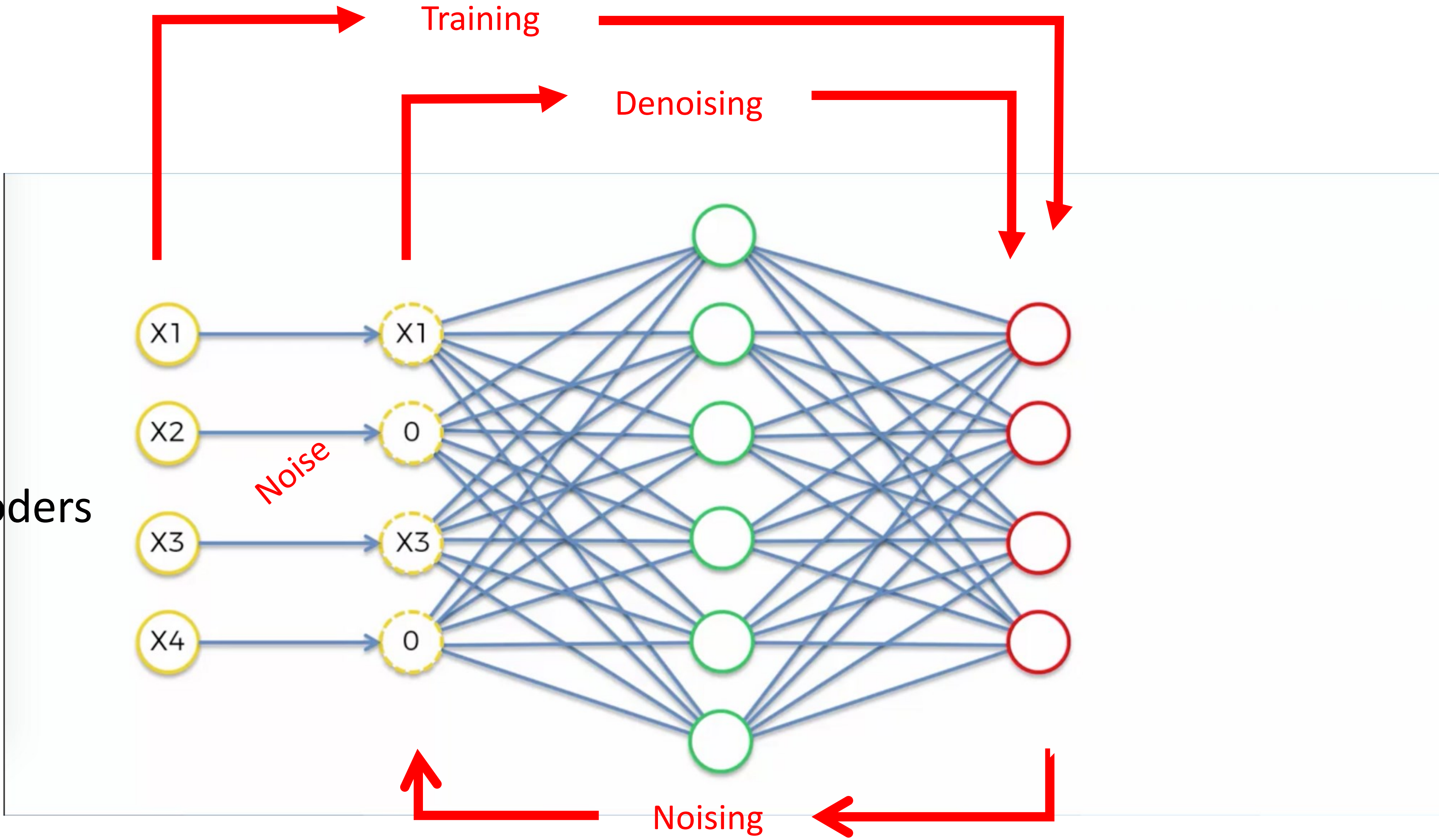


Challenge IV: corrupted data

[Aridor et.al 2003, Yadgar et.al 2015]

- Problems:
  - Missing values
  - Errors
  - Arbitrary noise

- Generative models
  - denoising autoencoders
  - (special type of variational autoencoders)



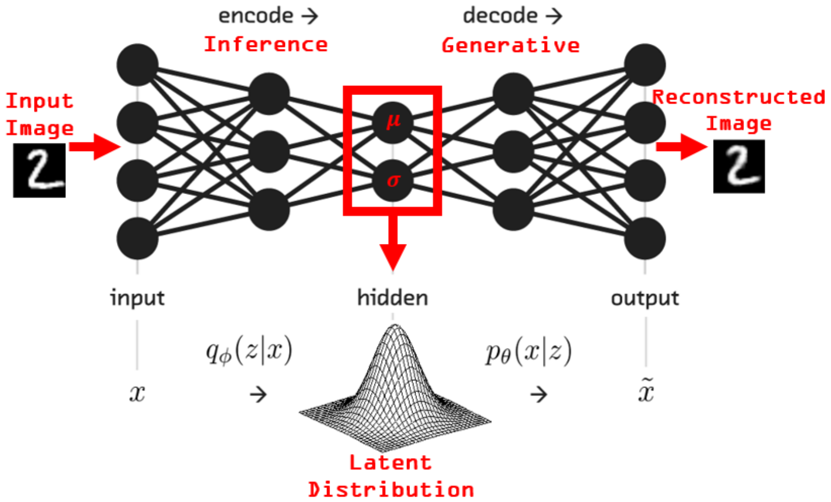
Challenge V: incorrect labels

[Palatin et.al 2008, Gabel et.al 2012]

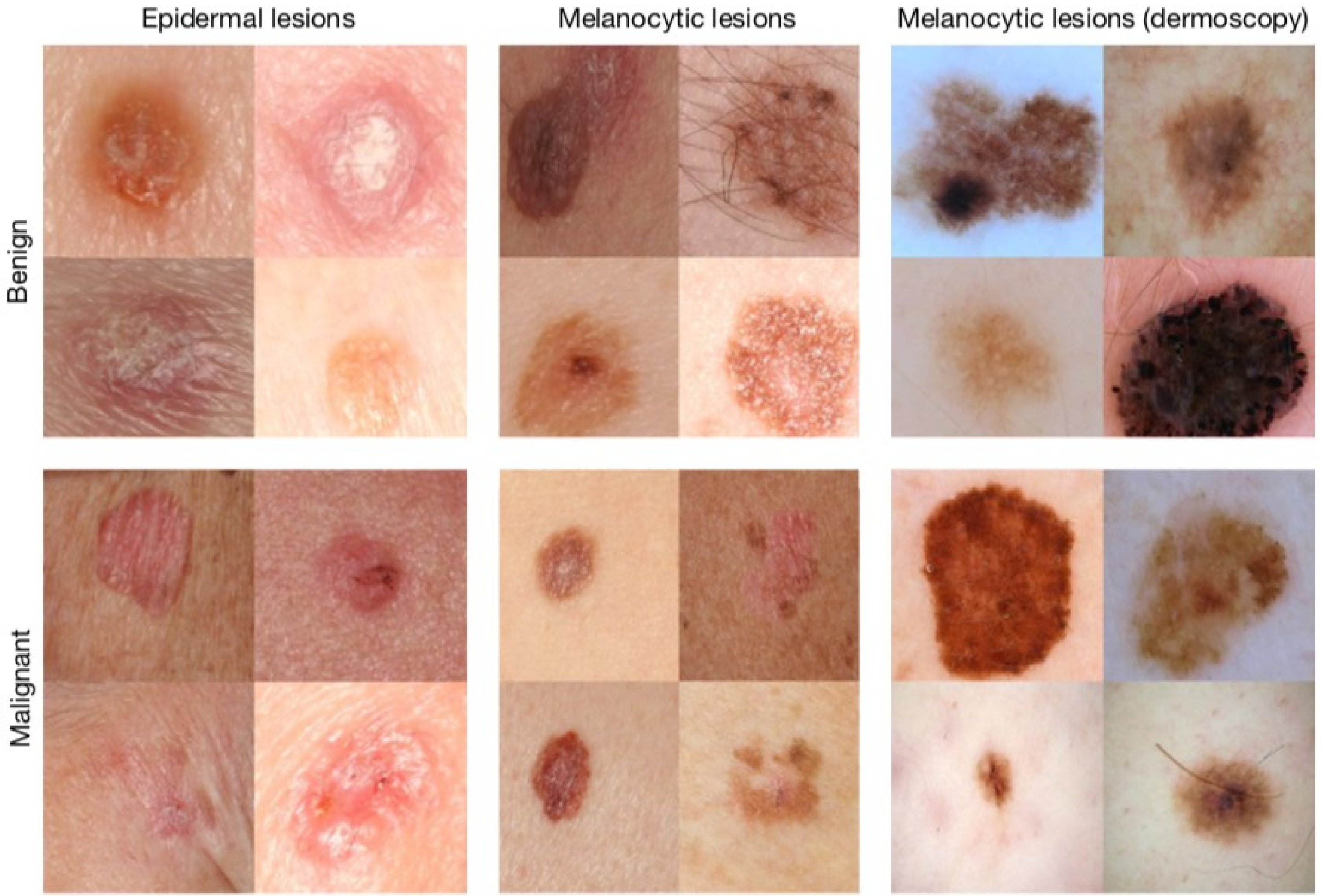
- Common solution – ***more data***
  - Expensive ***manual labor*** of experts
  - Assumption – most tagging are correct

- May sometimes apply ***clustering*** and then ***labeling of clusters***
  - Need to know the ***# of clusters***
  - Uses human-crafted ***features***.

- Popular for non-medical data – ***generative models***
  - variational autoencoders
  - Open – is this solution suitable for the ***medical domain***?



Clustering for Labels



example images from two disease classes. These test images highlight the difficulty of malignant versus benign discernment for the three medically critical classification tasks we consider: epidermal lesions, melanocytic lesions and melanocytic lesions visualized with a dermoscope. Example images reprinted with permission from the Edinburgh Dermofit Library (https://licensing.eri.ed.ac.uk/i/software/dermofit-image-library.html).

Dermatologist-level classification of skin cancer with deep neural networks

Andre Esteva[1]*, Brett Kuprel[1]*, Roberto A. Novoa[2,3], Justin Ko[2], Susan M. Swetter[2,4], Helen M. Blau[5] & Sebastian Thrun[6]

Challenge VI: non-symmetric classes

[Keren et.al 2006+2008+2018, Friedman et.al 2014]

Problems:

1. Not enough labeled examples for every class
2. Variation in # of labeled examples per class

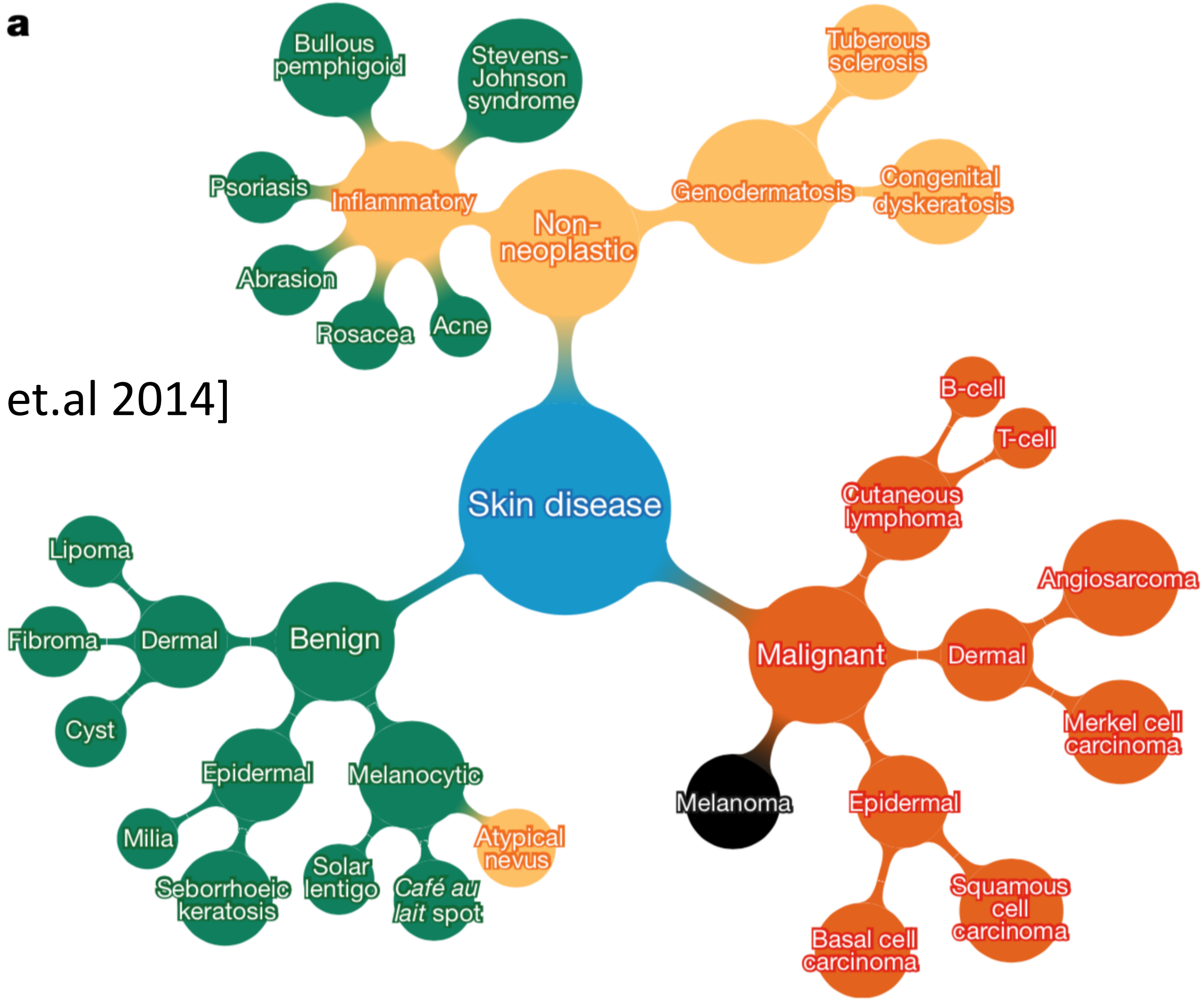


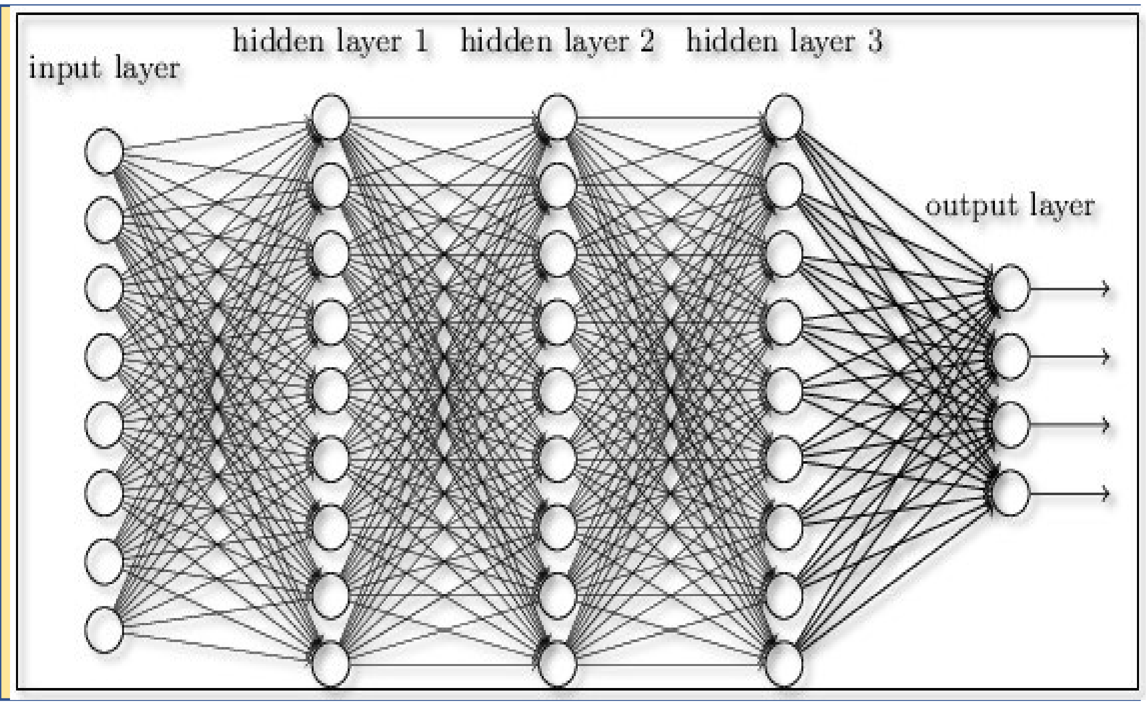
Figure 2 | A schematic illustration of the taxonomy and example test set images. a, A subset of the top of the tree-structured taxonomy of skin disease. The full taxonomy contains 2,032 diseases and is organized based on visual and clinical similarity of diseases. Red indicates malignant,

Dermatologist-level classification of skin cancer with deep neural networks

Andre Esteva[1*], Brett Kuprel[1*], Roberto A. Novoa[2,3], Justin Ko[2], Susan M. Swetter[2,4], Helen M. Blau[5] & Sebastian Thrun[6]

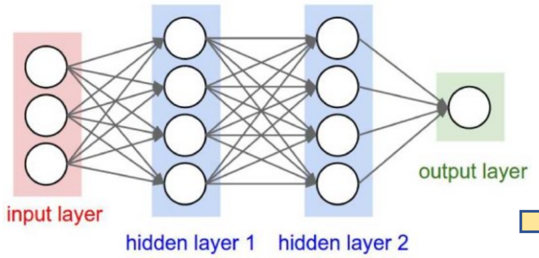
Challenge VII: Model Compaction

[Silberstein et.al 2008, Gabel et.al 2014]



Execute continuously?!

Train



Execute
(Accuracy?)

- “Teacher-student”:
  1. train a complex model
  2. let it train a light model

- “Quantization”:
  1. Edge weights can be quantized
  2. In fact, they can be rounded to be 0-1

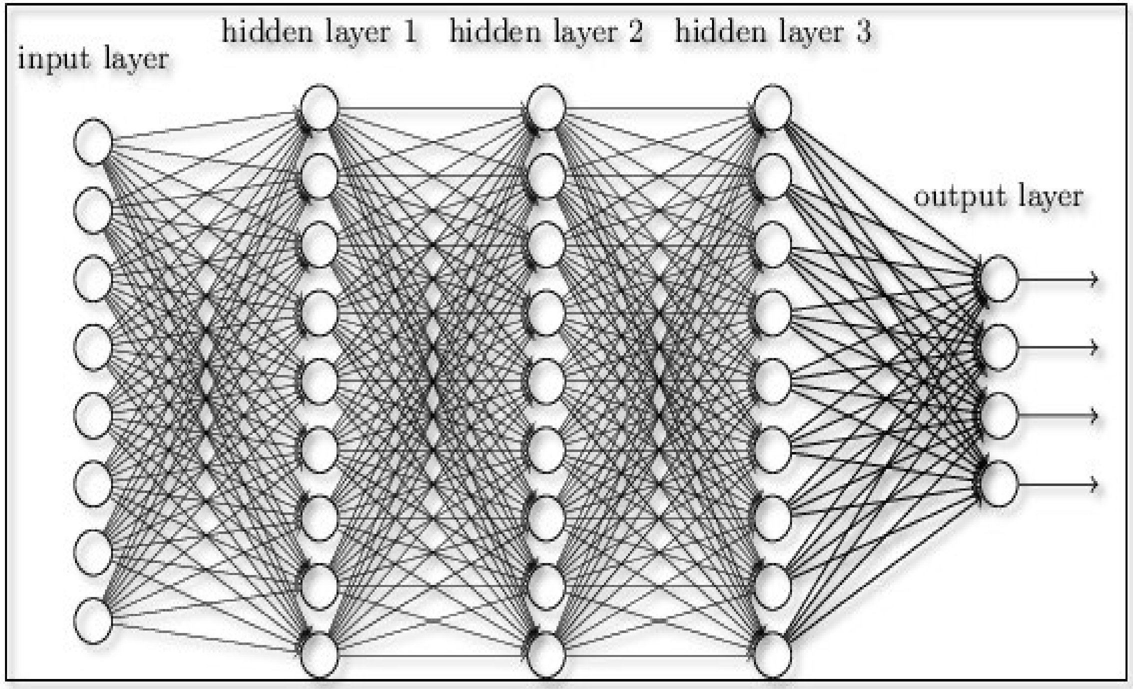
- Recent ideas:

Replace with memory tables

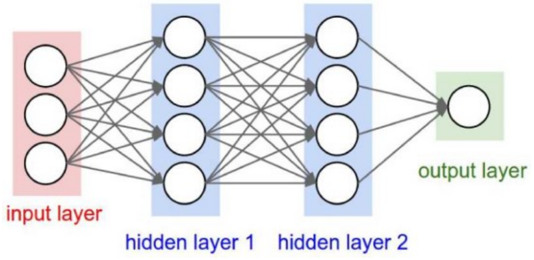
Challenge VIII: on-the-fly analytics

[Schuster et.al 2001, Wolff et.al 2005, Sharfman et.al 2008, Keren et.al 2012, Friedman et.al 2014, Kolchinsky et.al 2018]

data

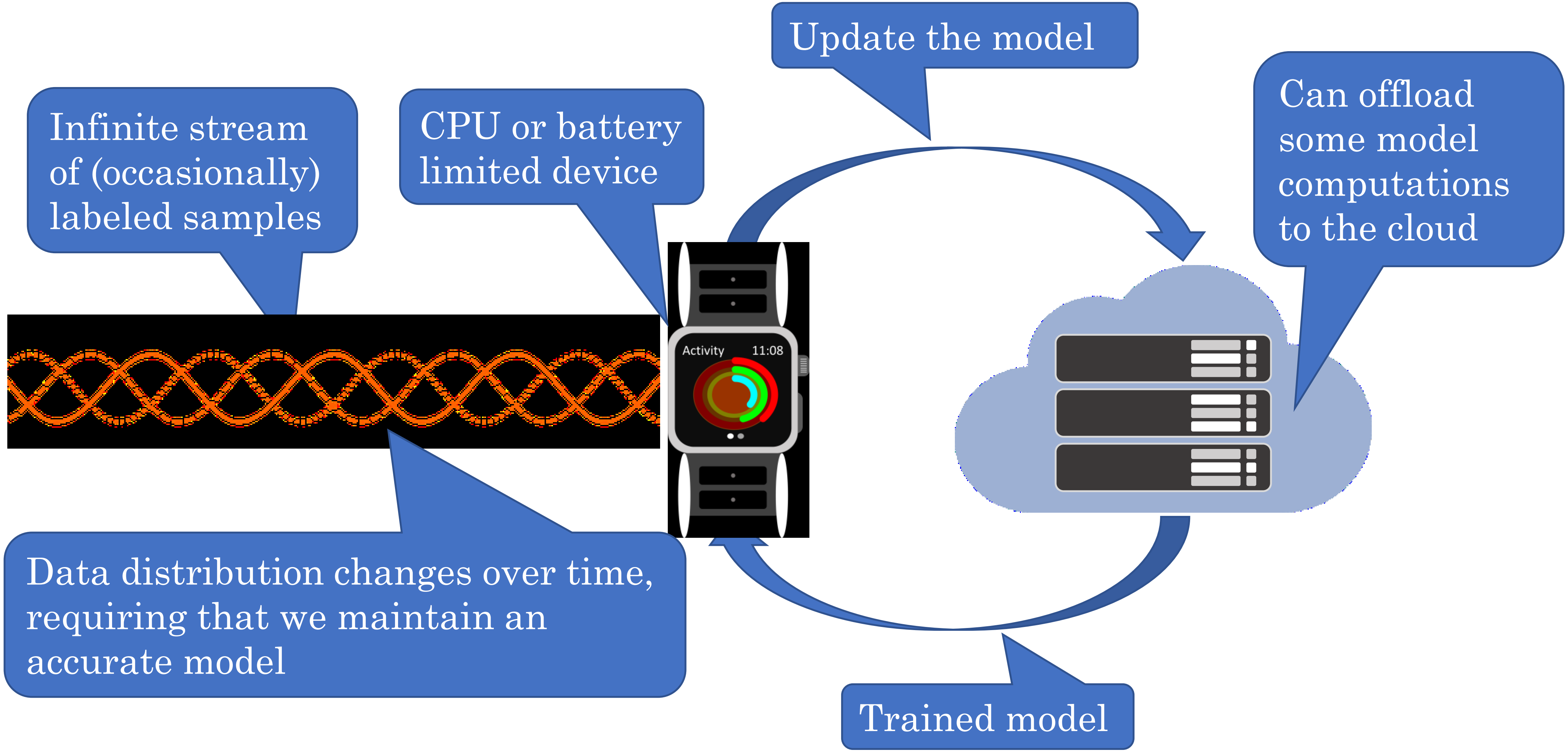


models



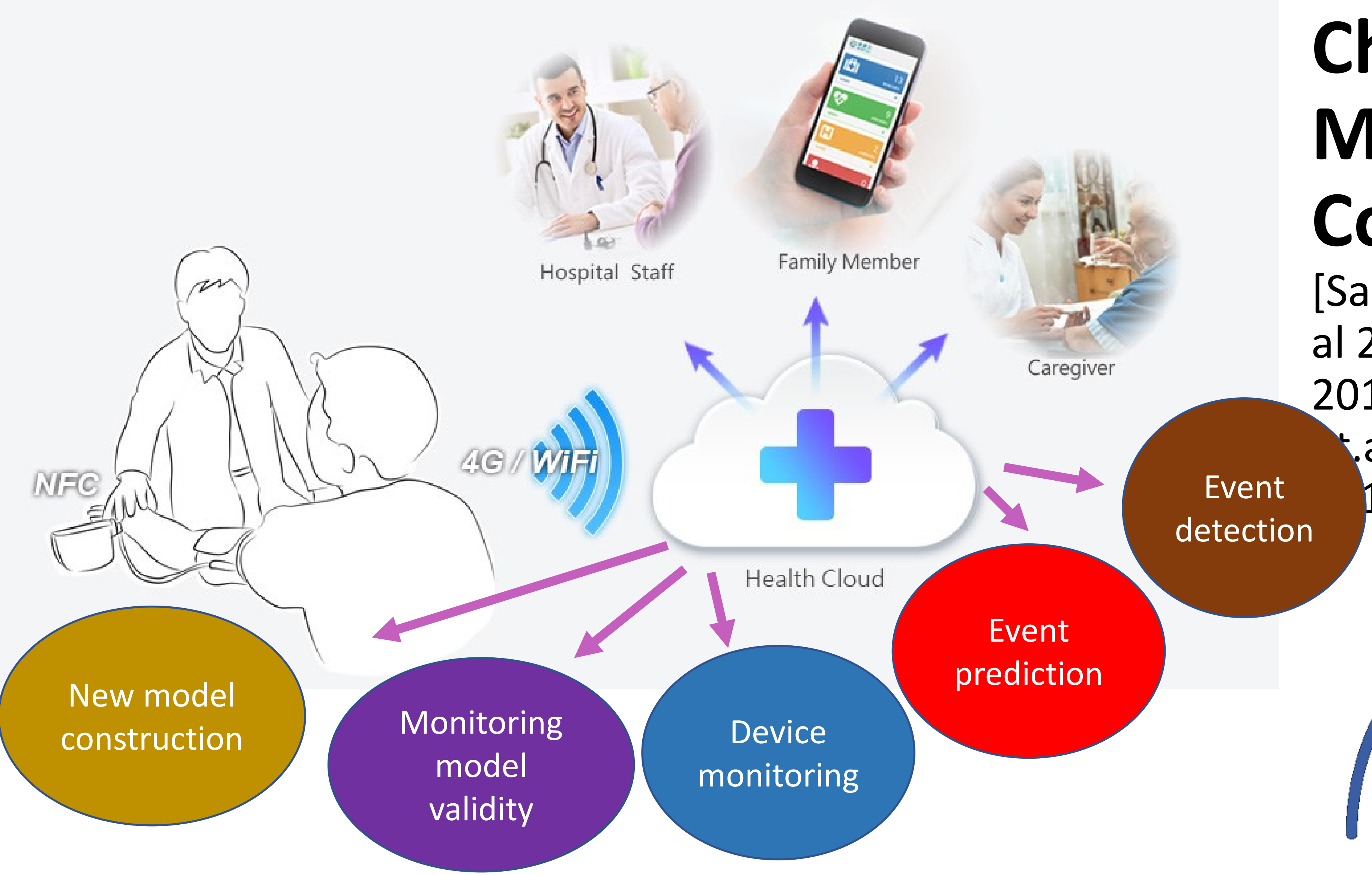
No false negatives!!!
Not too many false positives

Challenge IX: change-point detection



Challenge X: Minimizing Communication

[Sagy et.al 2010, Verner et. al 2011+2012, Keren et.al 2016+2014+2016, Lazerson et.al 2015, Friedman et.al 2018]



- What is “important” data?
- The concept of a “safe zone”

Challenge XI: Cyber Security

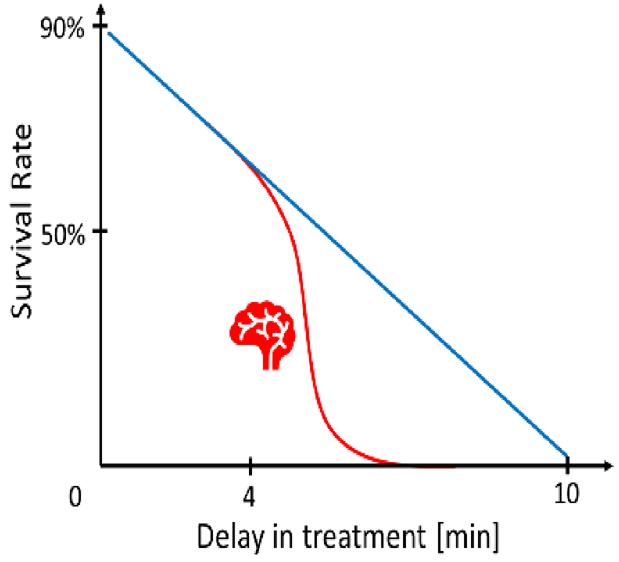
Data:

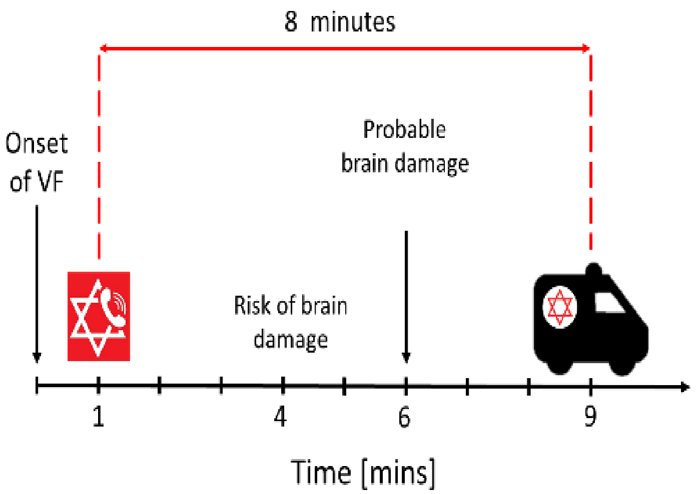
- Privacy preserving; anonymization
- Standards for medical data (GDPR, HIPAA,…).

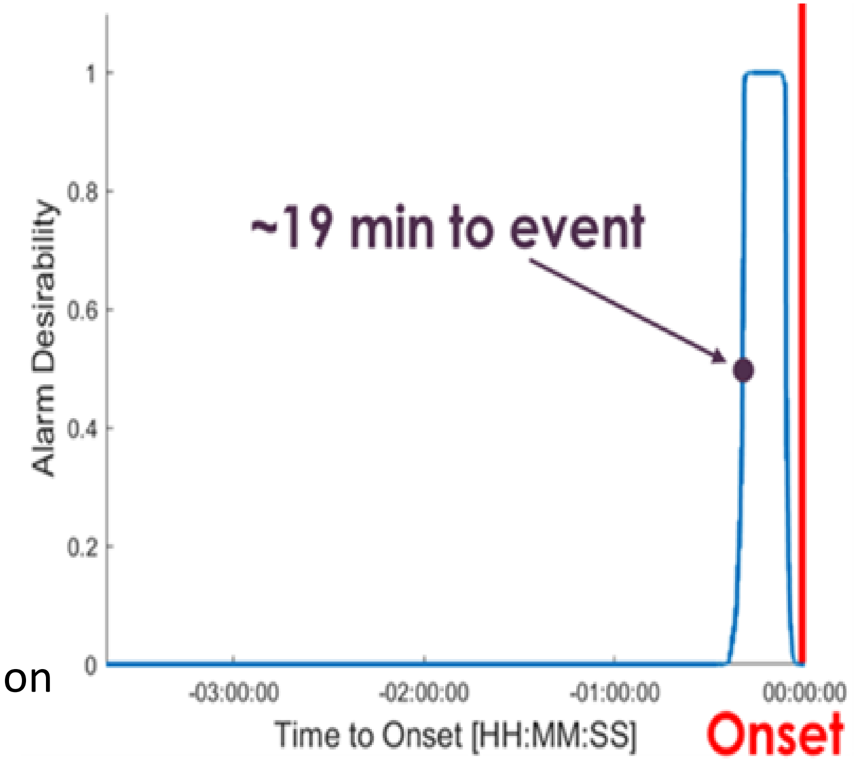
IoT security:

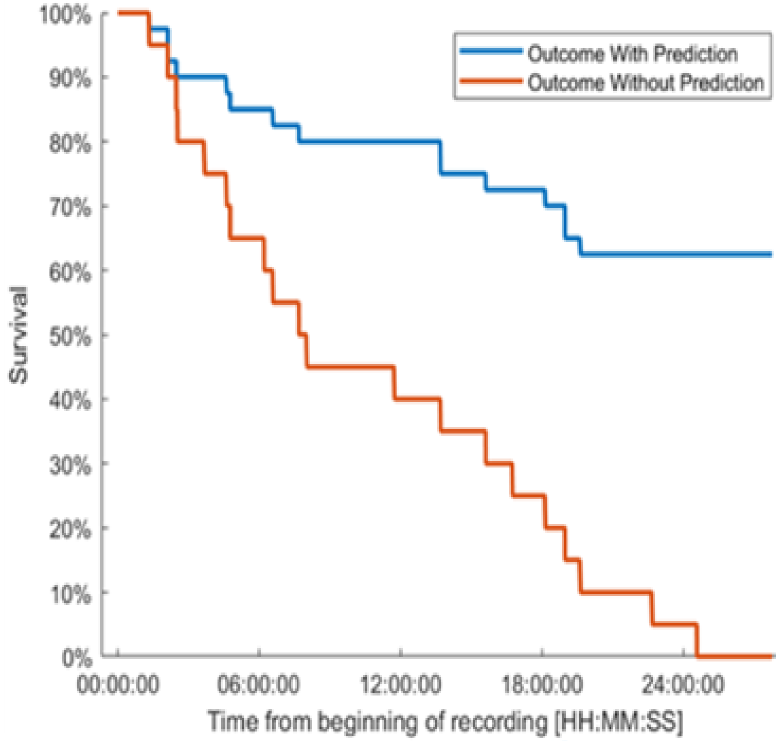
- Profiling devices in the wild
- Anomaly detection; pattern detection
- Imposing structure over million devices
- Enabling analytics over extremely large distributed systems

Challenge XII: short-term predictions. Predict VF?





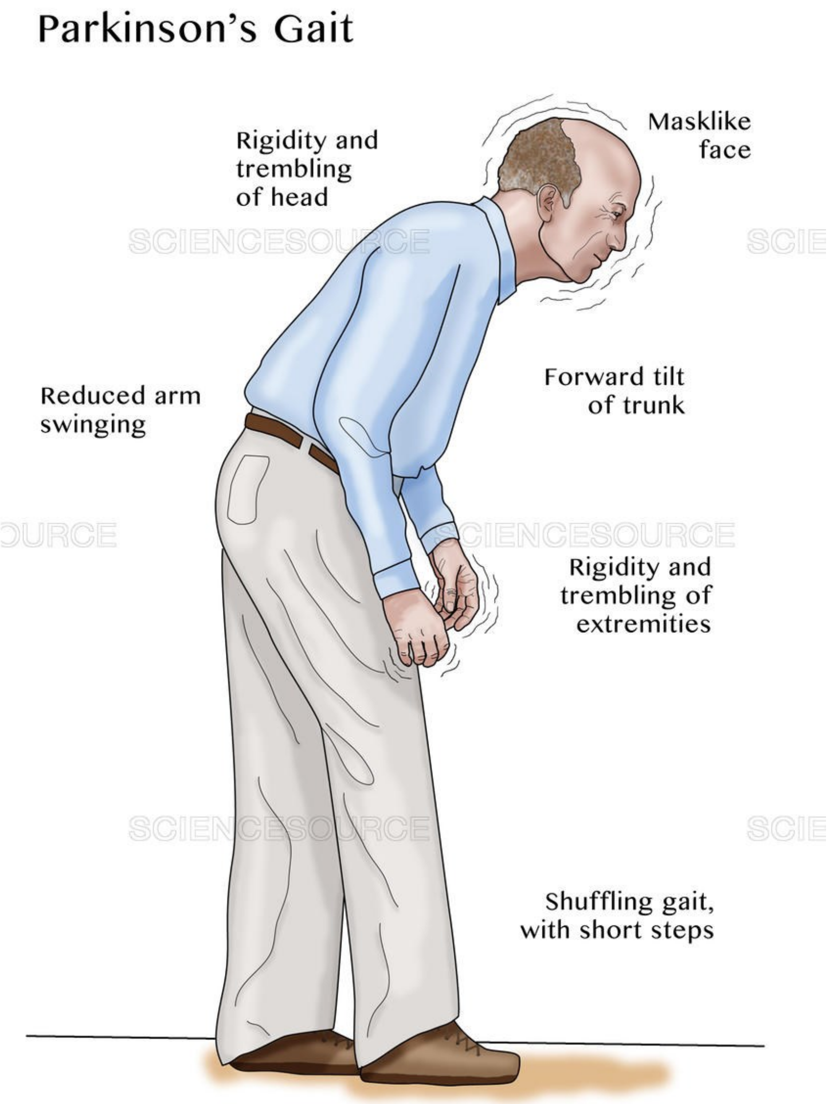




Best students project 2018
BioMedical Engineering, Technion
Edelstein & Keidar @Yaniv lab

Challenge XIII: long-term trend predictions

- How will this condition develop?
- How will average behavior look in 10 years from now?
- How will it look like when *this* medicine is given?
- ….. When this exercise routine is practiced?
- Etc.



My team @Technion

Learning:
- Neural Networks
- Transfer learning
- Monitoring, anomaly detection
- SGD acceleration

Scalable systems:
- Distributed computing
- Parallel Computing

Data management & Systems:
- Scalablility
- Data streams
- Edge computing
- Communication minimization

Cyber Security:
- IoT security
- Privacy

Questions?