Missingness & Representation

Matthew McDermott, MIT
CSC2541HS Guest Lecture

Outline

1. Missingness
2. Representation

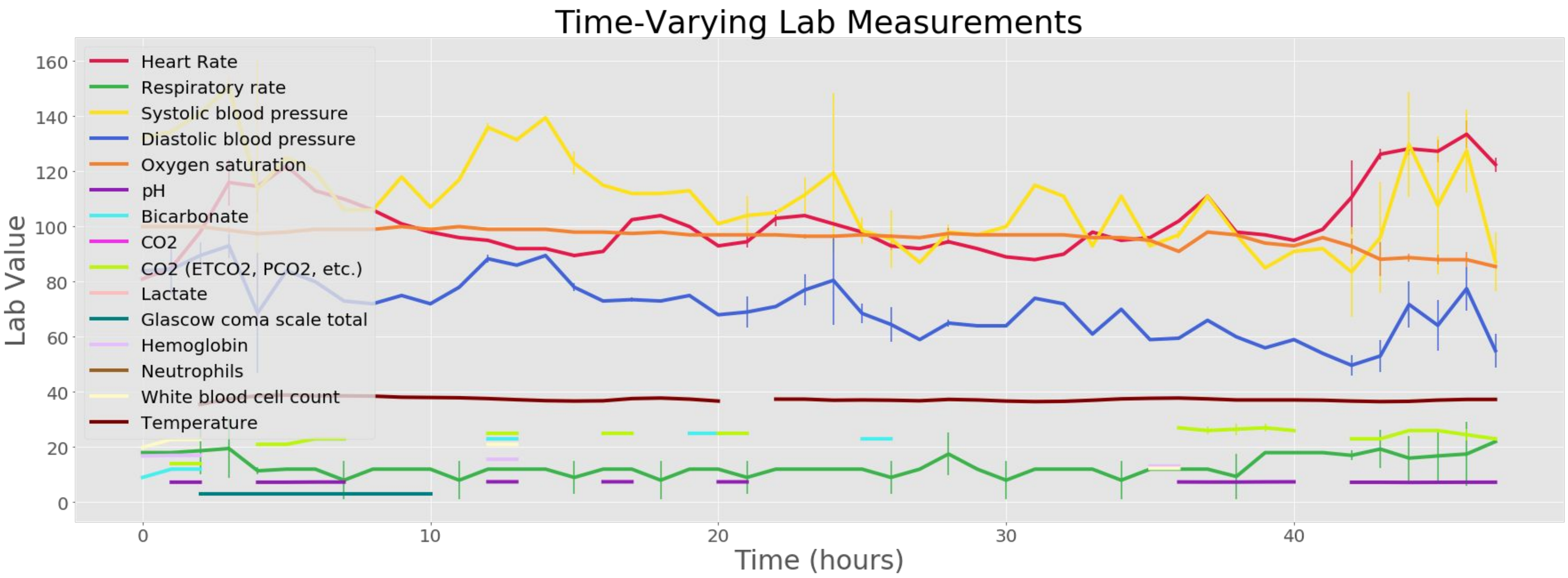
Outline

1. Missingness
    a. What is it?
    b. How does it affect us?
    c. What can we do?
2. Representation
    a. Why do we care?
    b. How can we find a good representation?
    c. How can we evaluate a representation?

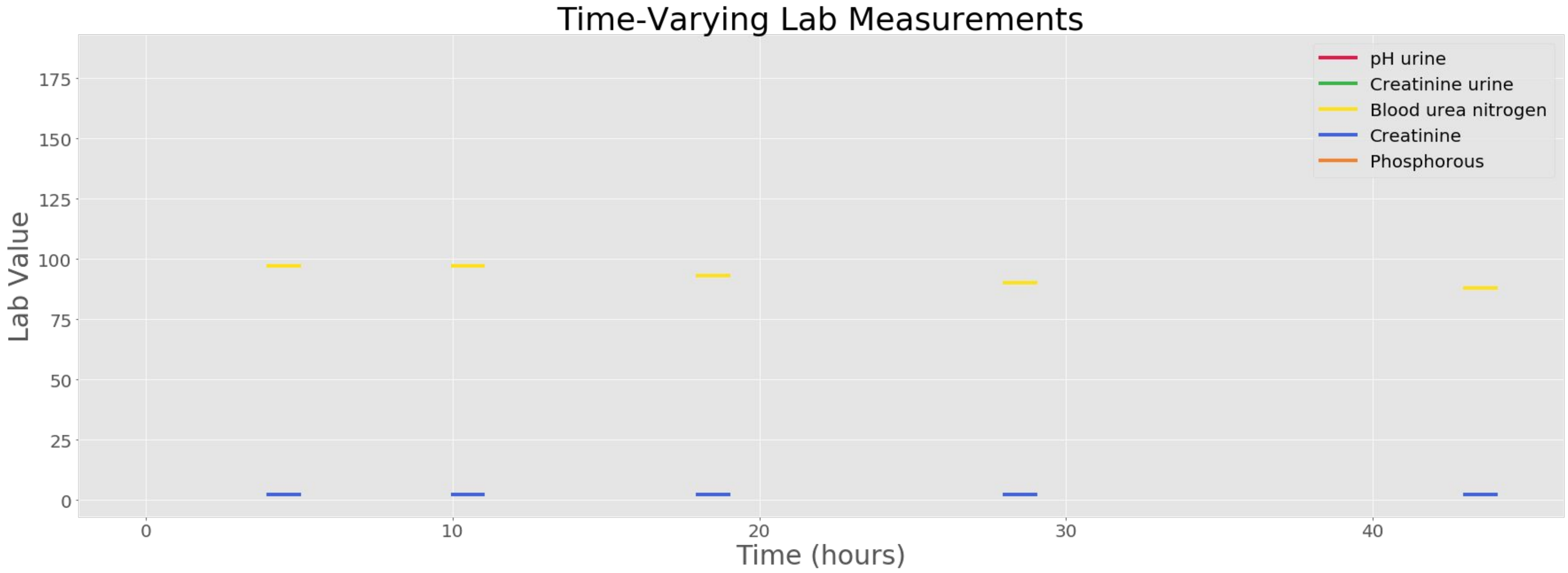
Outline

1. Missingness
   a. What is it?
   b. How does it affect us?
   c. What can we do?
2. Representation
   a. Why do we care?
   b. How can we find a good representation?
   c. How can we evaluate a representation? (Time Permitting)

What is Missingness?



Missing Data is Rampant in Healthcare Data



Different Kinds of Missingness (Data)

```
[**2173-2-6**] 10:02 PM
 CHEST (PORTABLE AP)                                           Clip # [**Clip Number (Radiology) 26360
**]
 Reason: R/O infiltrate, check ett position
 ______________________________________________________________________________
 [**Hospital 4**] MEDICAL CONDITION:
  35 year old man with AIDS, MS change, HONK, Fever
 REASON FOR THIS EXAMINATION:
  R/O infiltrate, check ett position
 ______________________________________________________________________________
                                 FINAL REPORT
 INDICATION: Fever s/p intubation.

 FINDINGS:

 The tip of the ETT is 6 cm above the carina. The feeding tube is noted
 coursing below the level of the diaphragm. Slight increased density is
 appreciated in the retrocardiac region and on this single shallow inspiration
```

Different Kinds of Missingness (Data)

```
[**2173-2-6**] 10:02 PM
 CHEST (PORTABLE AP)                                              Clip # [**Clip Number (Radiology) 26360
**]
 Reason: R/O infiltrate, check ett position
 ______________________________________________________________________________
 [**Hospital 4**] MEDICAL CONDITION:
  35 year old man with AIDS, MS change, HONK, Fever
 REASON FOR THIS EXAMINATION:
  R/O infiltrate, check ett position
 ______________________________________________________________________________
                                 FINAL REPORT
 INDICATION: Fever s/p intubation.

 FINDINGS:

 The tip of the ETT is 6 cm above the carina. The feeding tube
 coursing below the level of the diaphragm. Slight increased dens
 appreciated in the retrocardiac region and on this single shall
```

Radiological Image?

Why do we care?

- Missing data is inconvenient.

Why do we care?

- Missing data is inconvenient.
- Missing data is not going away.

Why do we care?

- Missing data is inconvenient.
- Missing data is not going away.
- Missing data is informative.

Why do we care?

- Missing data is inconvenient.
- Missing data is not going away.
- Missing data is informative (or confounding).

Missing Data Details

Data can be missing according to several regimes:

- Missing completely at random (MCAR)
- Missing at random (MAR)
- Missing not at random (MNAR)

Missing Data Details

Data can be missing according to several regimes:

- Missing completely at random (MCAR)
  - The observed pattern of missingness is independent from the observed or missing values.
- Missing at random (MAR)
- Missing not at random (MNAR)

Missing Data Details

Data can be missing according to several regimes:

- Missing completely at random (MCAR)
    - The observed pattern of missingness is independent from the observed or missing values.
- Missing at random (MAR)
    - The observed pattern of missingness is independent from the missing values (but may depend on the observed values).
- Missing not at random (MNAR)

Missing Data Details

Data can be missing according to several regimes:

- Missing completely at random (MCAR)
  - The observed pattern of missingness is independent from the observed or missing values.
- Missing at random (MAR)
  - The observed pattern of missingness is independent from the missing values (but may depend on the observed values).
- Missing not at random (MNAR)
  - All bets are off.

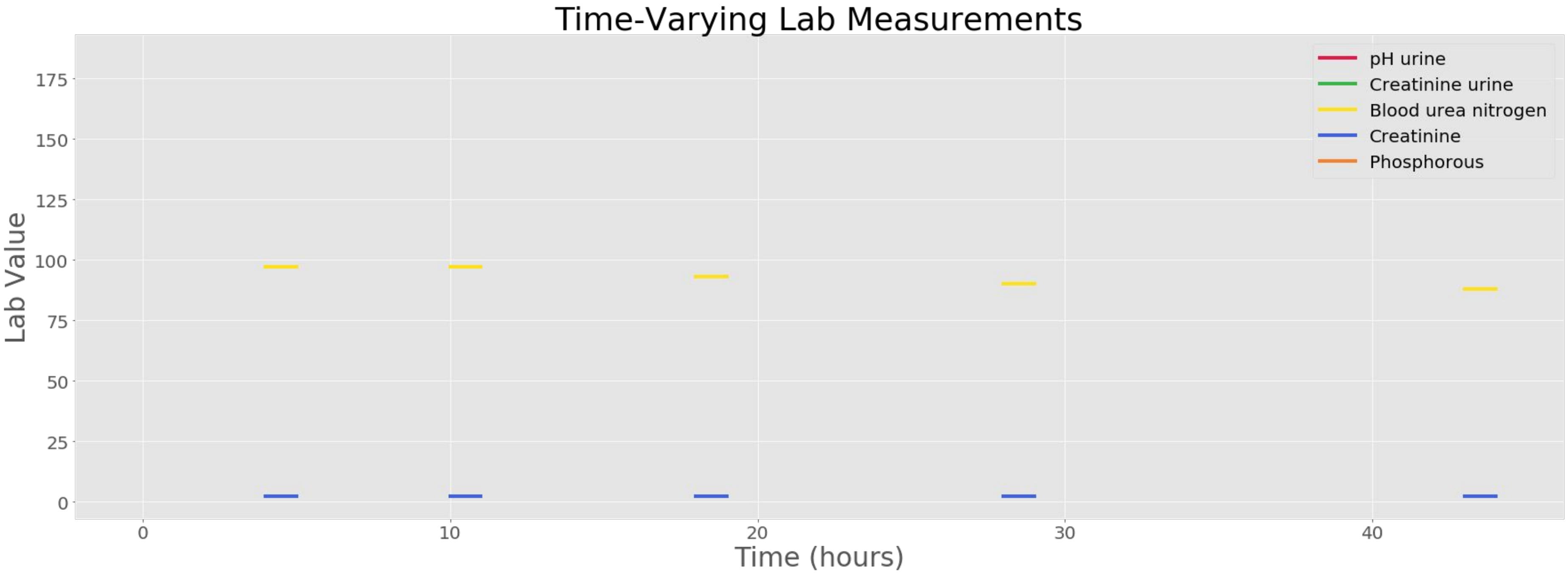
Missing Data Details

Data can be missing according to several regimes:

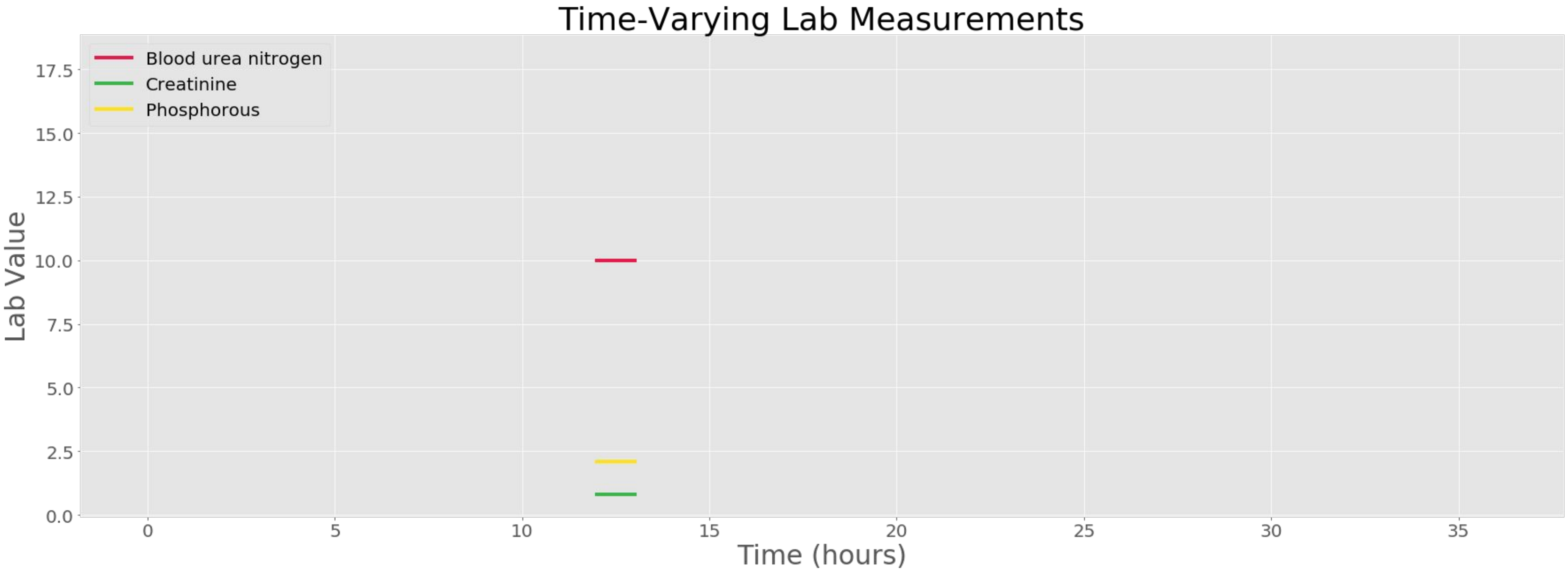
- Missing completely at random (MCAR)
  - The observed pattern of missingness is independent from the observed or missing values.
- Missing at random (MAR)
  - The observed pattern of missingness is independent from the missing values (but may depend on the observed values).
- Missing not at random (MNAR)
  - All bets are off.

Healthcare lives here.

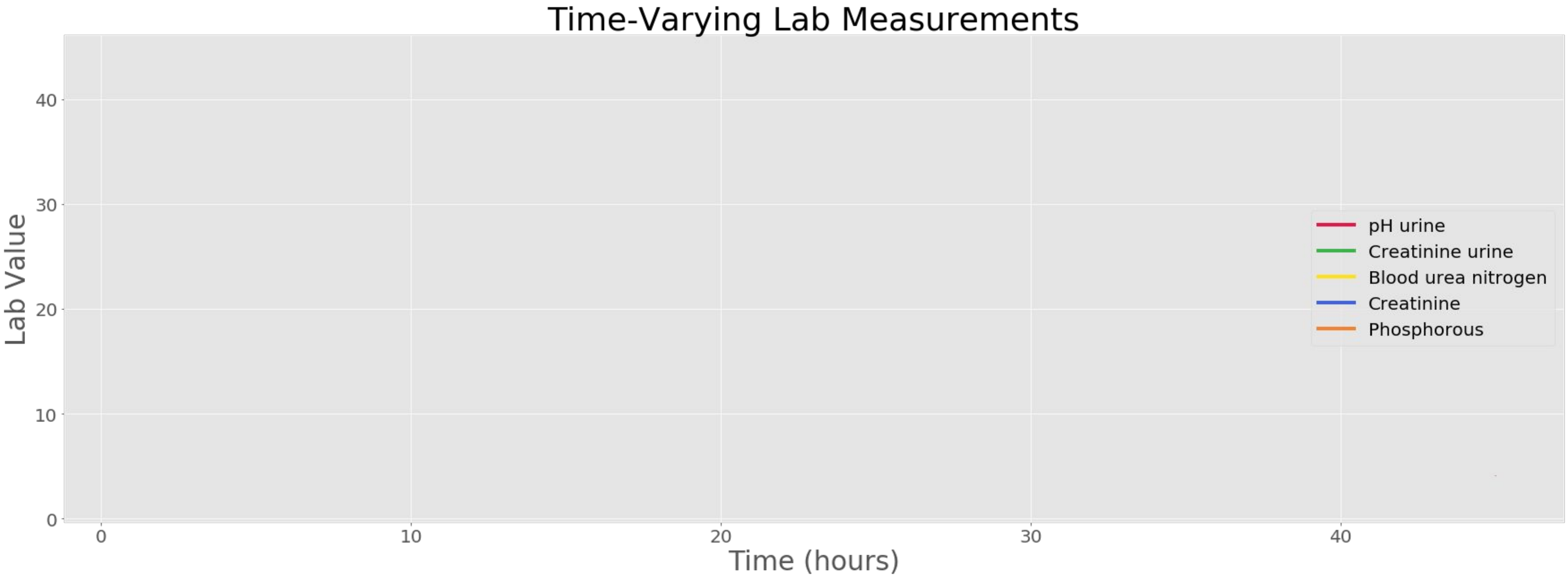
Missing Data is Information (Kidney)



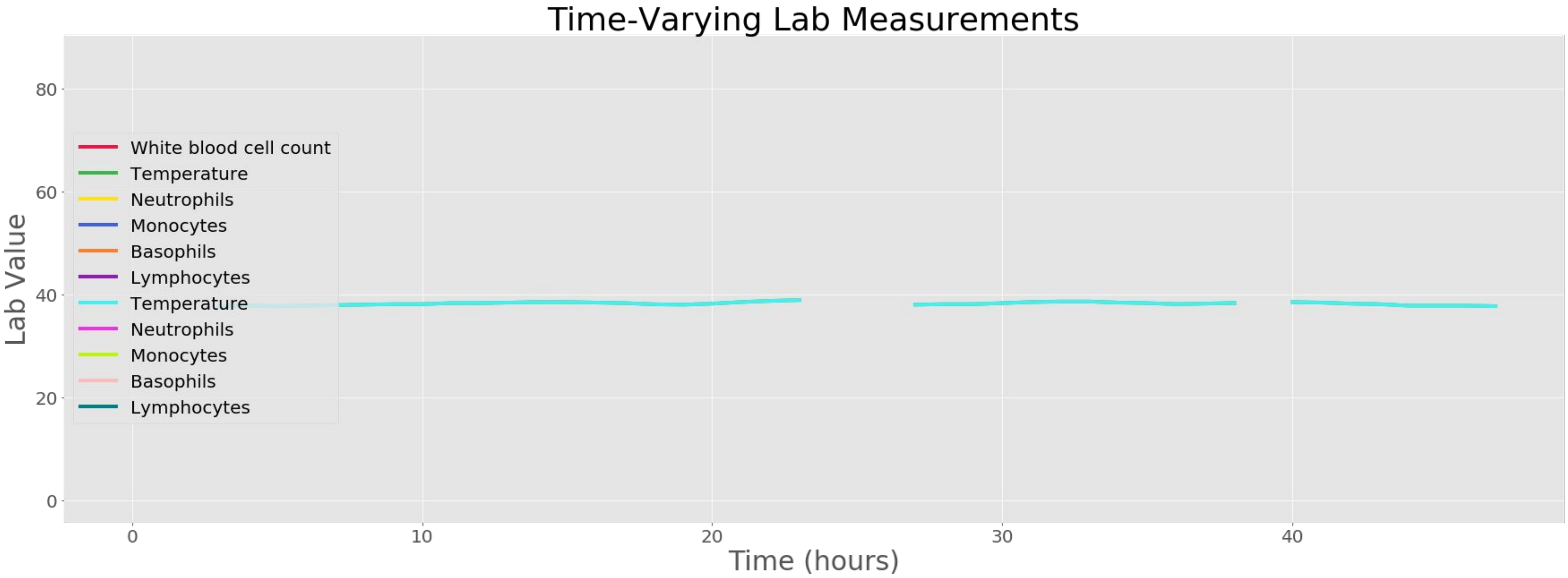
Missing Data is Information (Kidney)



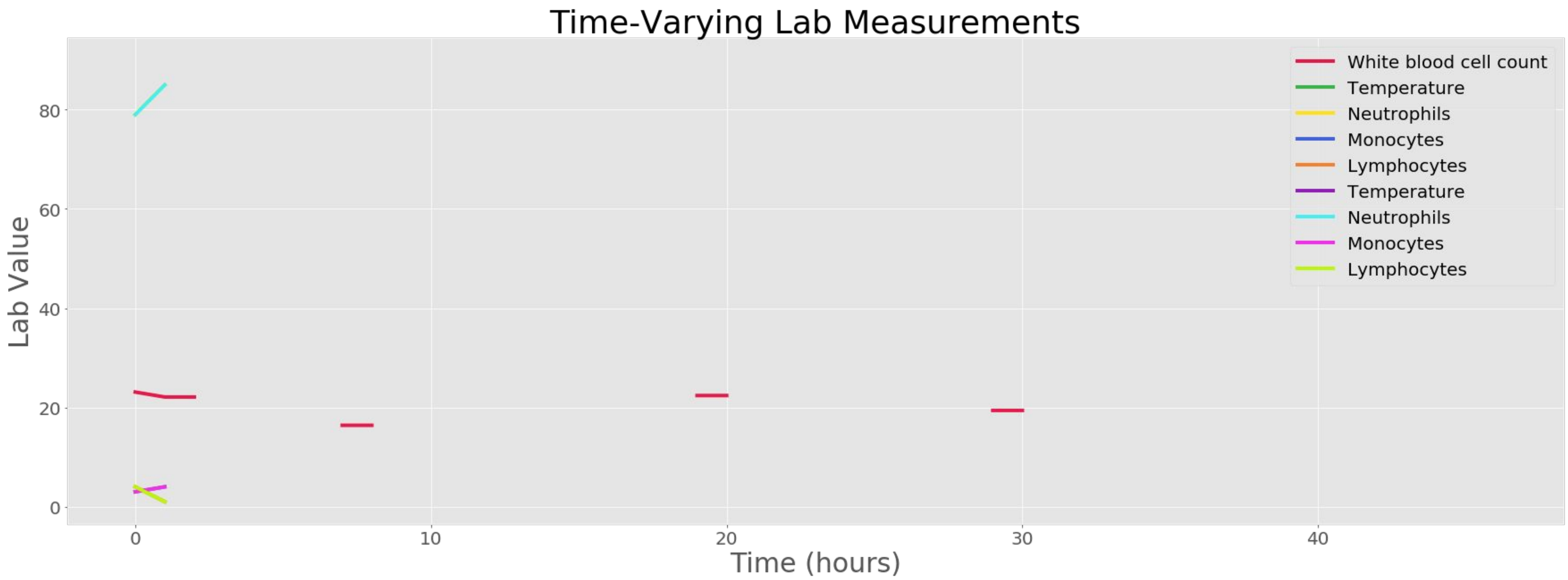
Missing Data is Information (Kidney)



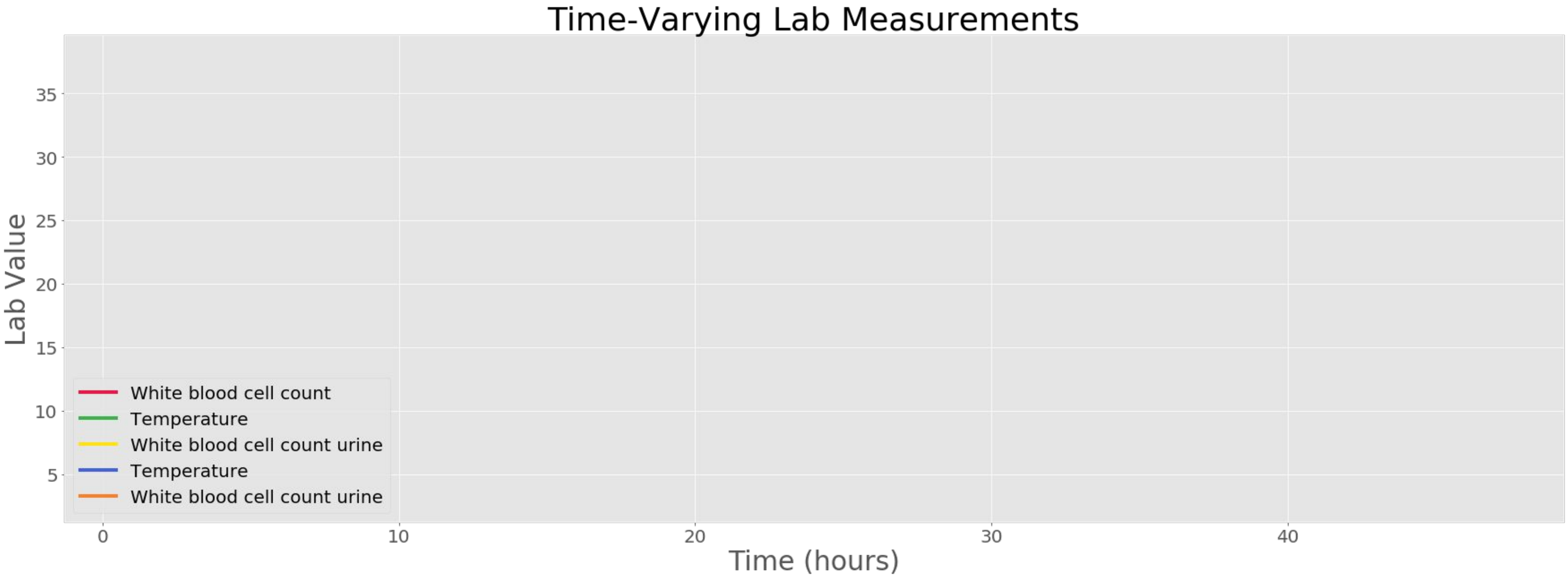
Missing Data is Information (Infection)



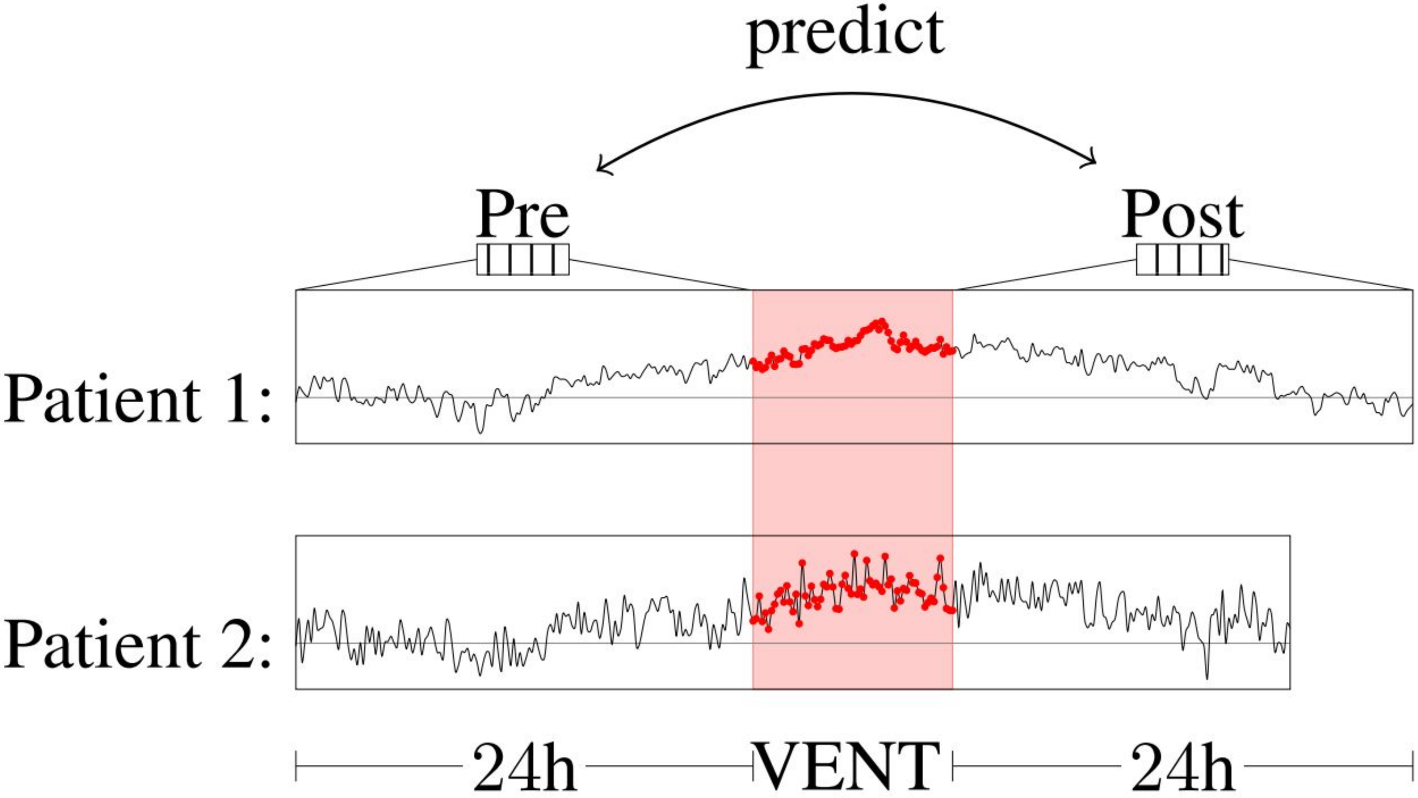
Missing Data is Information (Infection)



Missing Data is Information (Infection)



Missing Data is Confounding



How do we handle missing data?

RECURRENT NEURAL NETWORKS FOR MULTIVARIATE TIME SERIES WITH MISSING VALUES

Zhengping Che, Sanjay Purushot
Department of Computer Science
University of Southern California
Los Angeles, CA 90089, USA

Sparse Multi-Output Gaussian Processes for Medical Time Series Prediction

LIFANGC@PRINCETON.EDU

Modeling Irregularly Sampled Clinic

Satya Narayan Shukla, Benjamin M. Marl
College of Information and Computer Scienc
University of Massachusetts Amherst
Amherst, MA 01003
{snshukla,marlin}@cs.umass.edu

Modeling Missing Data in Clinical Time Series with RNNs

Zachary C. Lipton ZLIPTON@CS.UCSD.EDU
Department of Computer Science and Engineering
University of California, San Diego
La Jolla, CA 92093, USA

David C. Kale KALE@ISI.EDU
USC Information Sciences Institute
Marina del Rey, CA, USA

Randall Wetzel RWETZEL@CHLA.USC.EDU
Laura P. and Leland K. Whittier Virtual Pediatric Intensive Care Unit
Children's Hospital LA
Los Angeles, CA 90089

Imputation

1. Statistical Timeseries Forecasting: ARMA/ARIMA/ARIMAX, etc.
2. Easy Baselines: Constant infilling, Sample & Hold (+ indicators), Interpolation
3. Traditional Imputation: MICE/3D-MICE, MissForest, Matrix/Tensor Completion
4. Gaussian Processes
5. Advanced neural methods (GRU-D, GANs, etc.)

Imputation

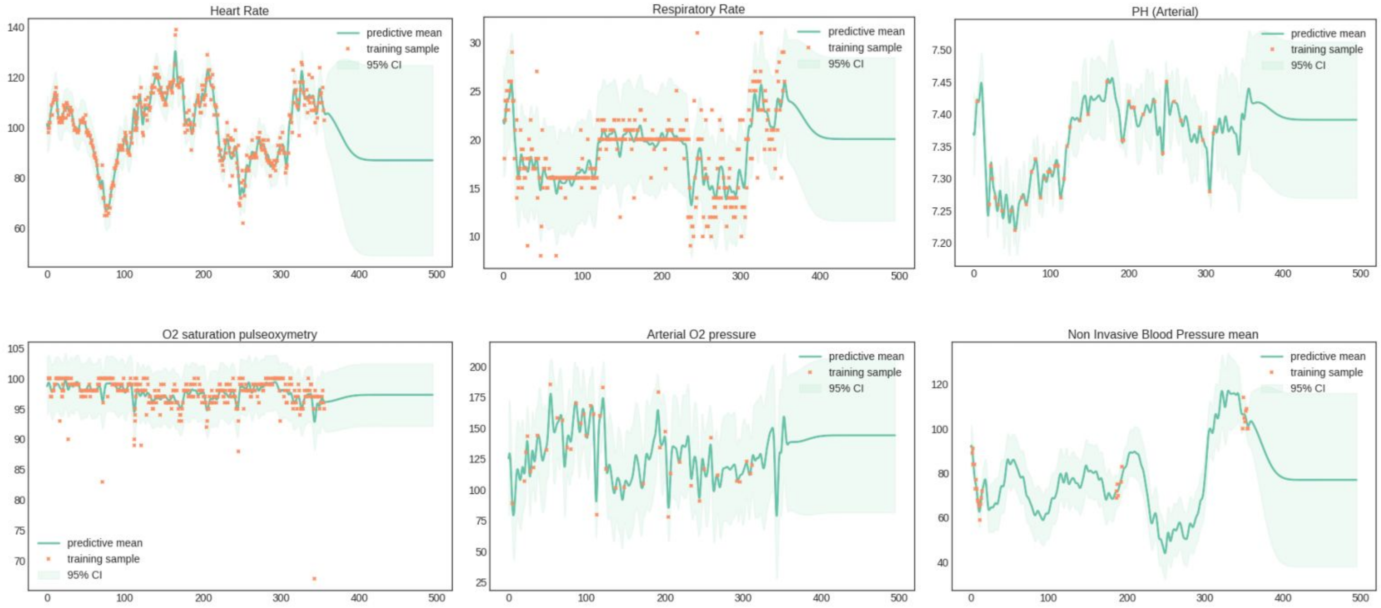


Figure 2: Example trajectories of six vital signs for a single admission, following imputation using Gaussian processes. Twelve vital signs are jointly modeled by the GP.

Prasad, Niranjani, et al. "A reinforcement learning approach to weaning of mechanical ventilation in intensive care units." *arXiv preprint arXiv:1704.06300* (2017).

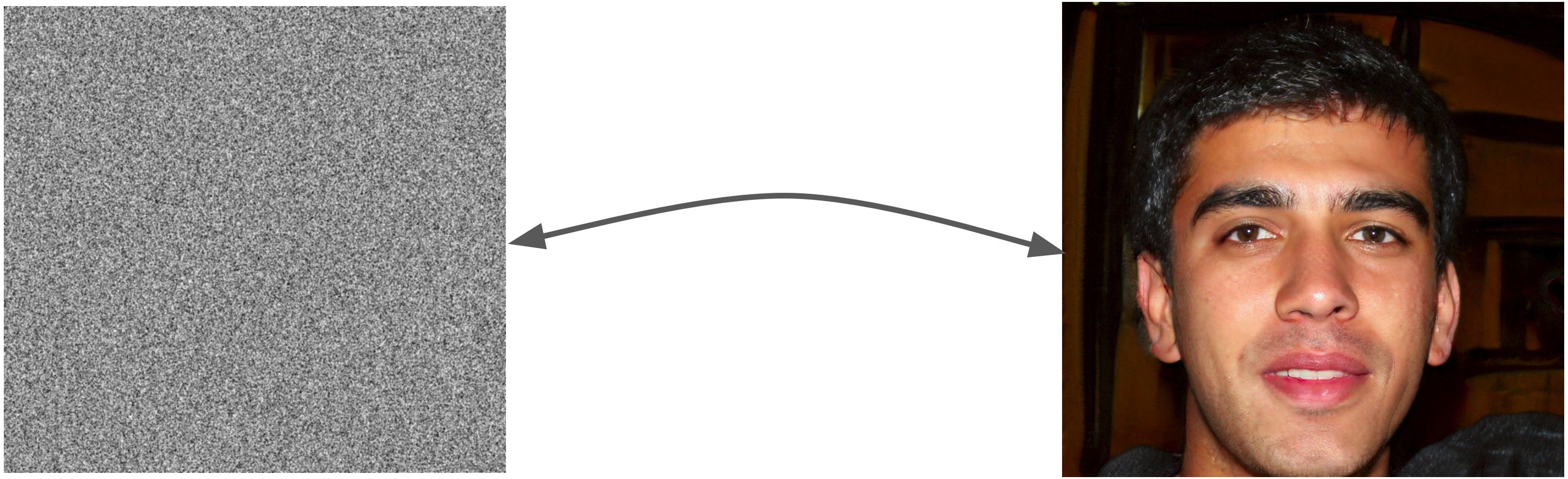
GANs for Imputation

GAIN: Missing Data Imputation using Generative Adversarial Nets

Jinsung Yoon [1,*] **James Jordon** [2,*] **Mihaela van der Schaar** [1,2,3]

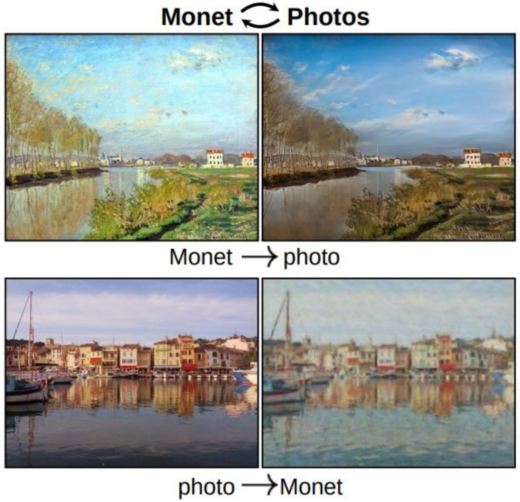
GANs for Imputation

GANs for Imputation



GANs for Imputation

Figure 6. Qualitative comparisons with Deepfillv1 [18] on the CelebA-HQ validation sets.



Left: Jo, Youngjoo, and Jongyoul Park. "SC-FEGAN: Face Editing Generative Adversarial Network with User's Sketch and Color." arXiv preprint arXiv:1902.06838 (2019).
Middle: Zhu, Jun-Yan, et al. "Unpaired image-to-image translation using cycle-consistent adversarial networks." Proceedings of the IEEE International Conference on Computer Vision. 2017.
Right: https://thispersondoesnotexist.com/

GAIN: Generative Adversarial Imputation

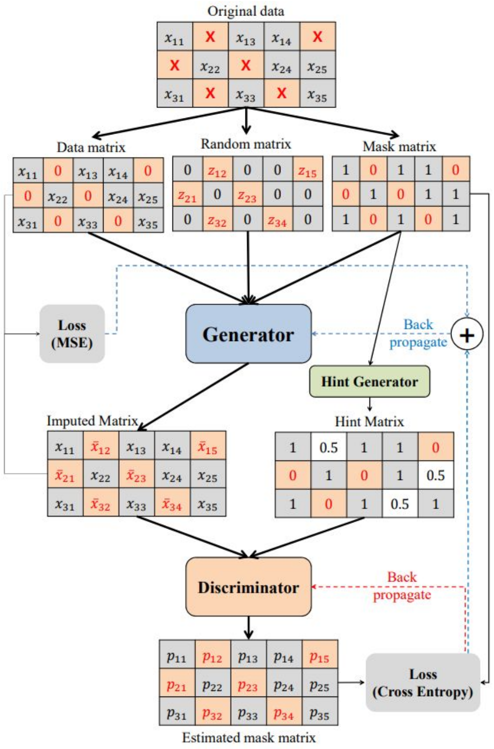


Figure 1. The architecture of GAIN

Imputation Papers

1. GAIN: https://arxiv.org/pdf/1806.02920.pdf
2. GRU-D: https://www.nature.com/articles/s41598-018-24271-9
3. GP Imputation: https://arxiv.org/pdf/1704.06300.pdf
4. Interpolation-prediction network: https://arxiv.org/pdf/1812.00531.pdf

Table 1: Performance on mortality and length of stay prediction tasks on MIMIC-III. Loss: Cross-Entropy Loss, MedAE: Median Absolute Error (in days), EV: Explained variance

Model	**Classification**			**Regression**	
	AUC	**AUPRC**	**Loss**	**MedAE**	**EV score**
Log/LinReg	0.772 ± 0.013	0.303 ± 0.018	0.240 ± 0.003	3.528 ± 0.072	0.043 ± 0.012
SVM	0.671 ± 0.005	0.300 ± 0.011	0.260 ± 0.002	3.523 ± 0.071	0.042 ± 0.011
AdaBoost	0.829 ± 0.007	0.345 ± 0.007	0.663 ± 0.000	4.517 ± 0.234	0.100 ± 0.012
RF	0.826 ± 0.008	0.356 ± 0.010	0.315 ± 0.025	3.113 ± 0.125	0.117 ± 0.035
GRU-M	0.831 ± 0.007	0.376 ± 0.022	0.220 ± 0.004	3.140 ± 0.196	0.131 ± 0.044
GRU-F	0.821 ± 0.007	0.360 ± 0.013	0.224 ± 0.003	3.064 ± 0.247	0.126 ± 0.025
GRU-S	0.843 ± 0.007	0.376 ± 0.014	0.218 ± 0.005	2.900 ± 0.129	0.161 ± 0.025
GRU-D	0.835 ± 0.013	0.359 ± 0.025	0.225 ± 0.009	$\mathbf{2.891 \pm 0.103}$	0.146 ± 0.051
Proposed	$\mathbf{0.853 \pm 0.007}$	$\mathbf{0.418 \pm 0.022}$	$\mathbf{0.210 \pm 0.004}$	$\mathbf{2.862 \pm 0.166}$	$\mathbf{0.245 \pm 0.019}$

Opportunities

1. Improved imputation methods. How do forecasting, GP, or adversarial methods compare to GRU-D/interpolation prediction network? Can we incorporate uncertainty offered by GPs usefully into downstream tasks? Can we make other models offer uncertainty?
2. Can we model the decision process by which clinicians choose what to measure and what to omit? How would this be helpful in downstream tasks? Can this help account for the MNAR nature of healthcare missingness?
3. Can we control for the confounding effects of missingness? Can we learn a model on underlying physiology from retrospective, care-byproduct data?

Representation

1. Missingness
    a. What is it?
    b. How does it affect us?
    c. What can we do?
2. **Representation**
    a. Why do we care?
    b. How can we find a good representation?
    c. How can we evaluate a representation?

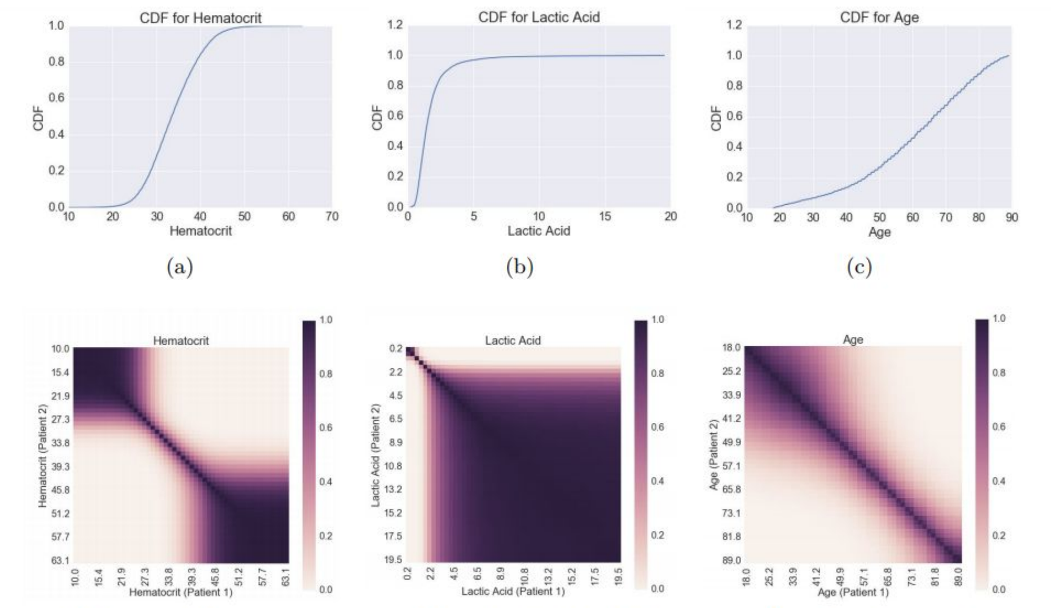
Representation: Why do we care?

Representations define a notion of "similarity"

Closer in "Conceptual Space"

Closer in "Pixel Space"

Representations learn a notion of similarity



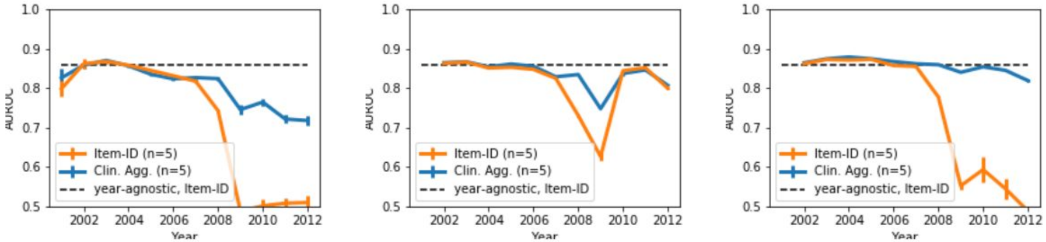
(a) (b) (c)

(d) Kernel on Hematocrit (e) Kernel on Lactic Acid (f) Kernel on Patient Age

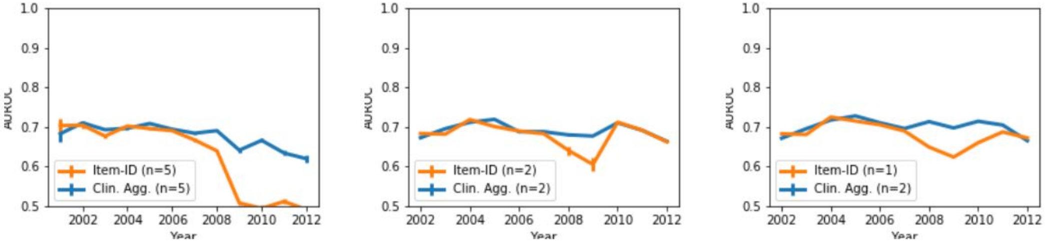
Figure 1: Examples of the kernel $k_{j,c}(x, z)$ in (1) with $c = 5$ on three features evaluated on adult ICU population: Hematocrit, Lactic Acid, and Patient Age

Conroy, Bryan, Minnan Xu-Wilson, and Asif Rahman. "Patient Similarity Using Population Statistics and Multiple Kernel Learning." Machine Learning for Healthcare Conference. 2017.

Representations can stabilize changing data



(a) Mortality AUC, models trained on 2001-2002 data. (b) Mortality AUC, models trained yearly on prior year only. (c) Mortality AUC, models trained yearly on all prior data.

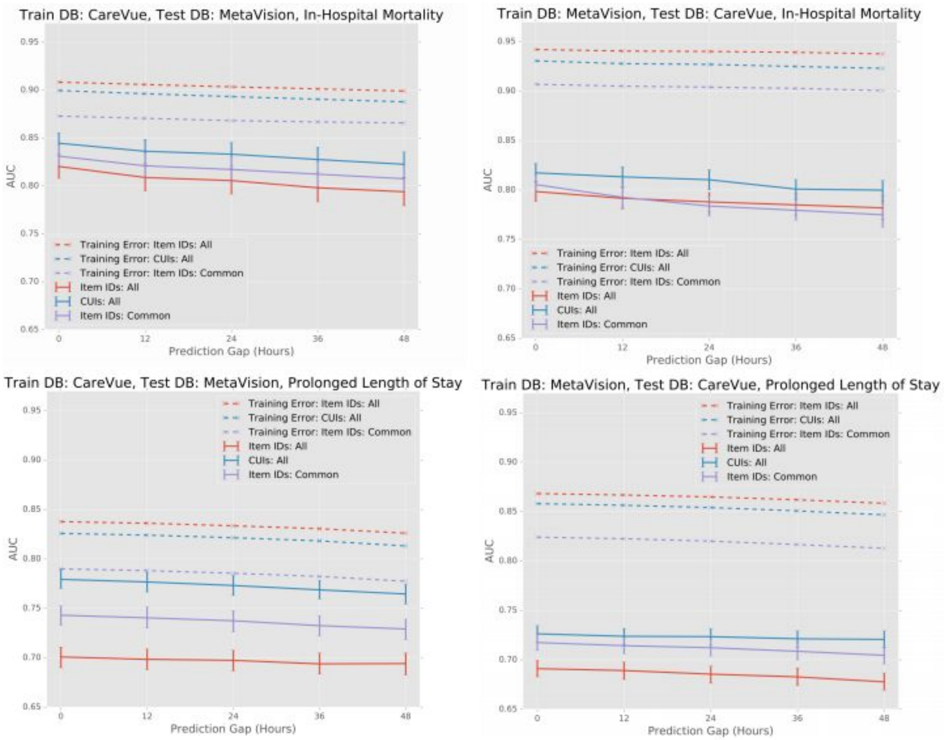


(d) LOS AUC, models trained on 2001-2002 data. (e) LOS AUC, models trained yearly on prior year only. (f) LOS AUC, models trained yearly on all prior data.

Figure 1: Performance of RF classifiers using `Item-Id` and `Clinically Aggregated` representations on mortality (top) and LOS prediction (bottom). Error bars indicate ± standard error.

Nestor, Bret, et al. "Rethinking clinical prediction: Why machine learning must consider year of care and feature aggregation." arXiv preprint arXiv:1811.12583 (2018).

Representations can stabilize changing data



Gong, Jen J., et al. "Predicting clinical outcomes across changing electronic health record systems." *Proceedings of the 23rd ACM SIGKDD International Conference on Knowledge Discovery and Data Mining*. ACM, 2017.

Representations can join disparate modalities

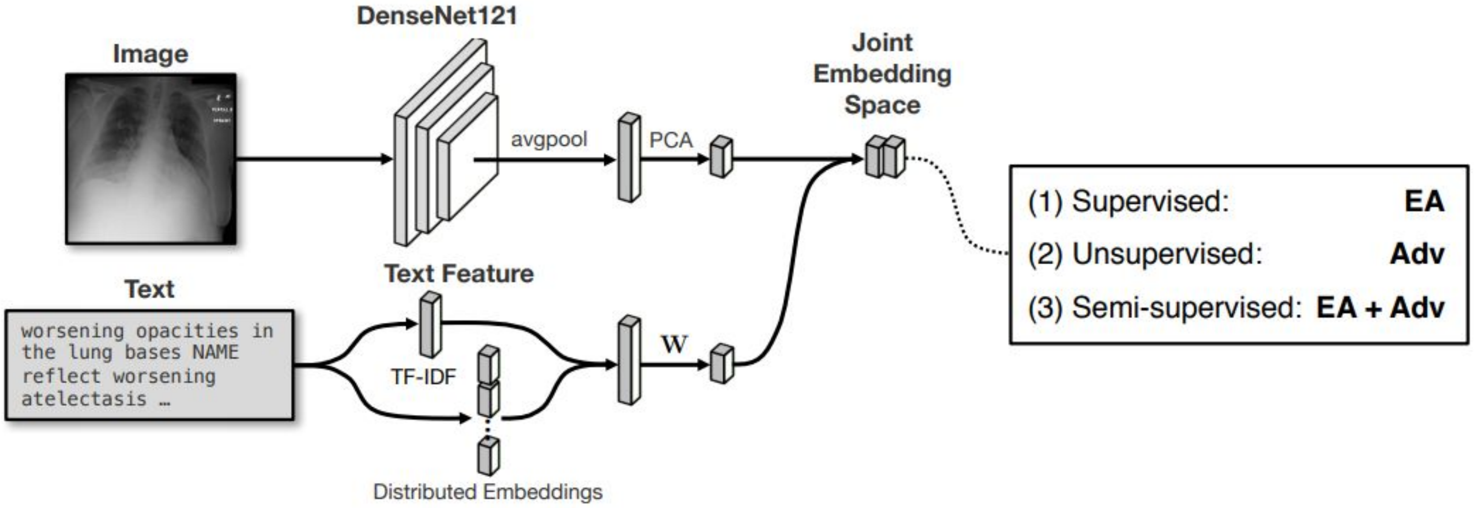
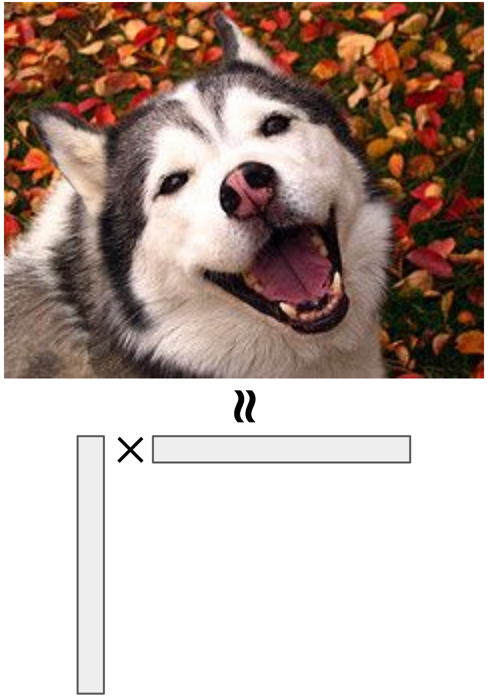
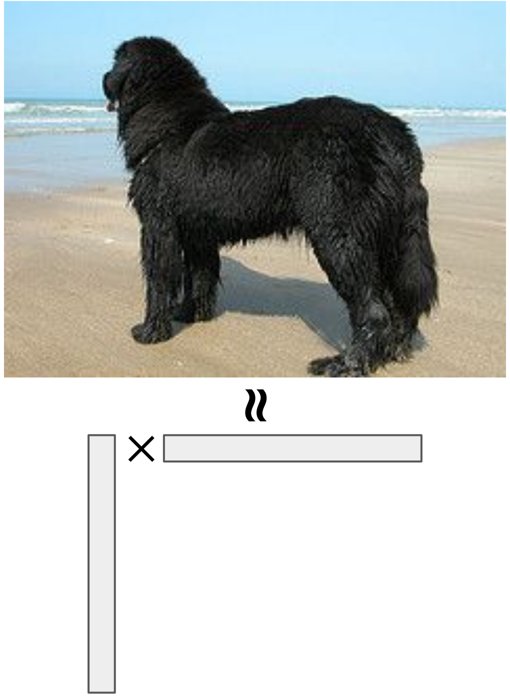


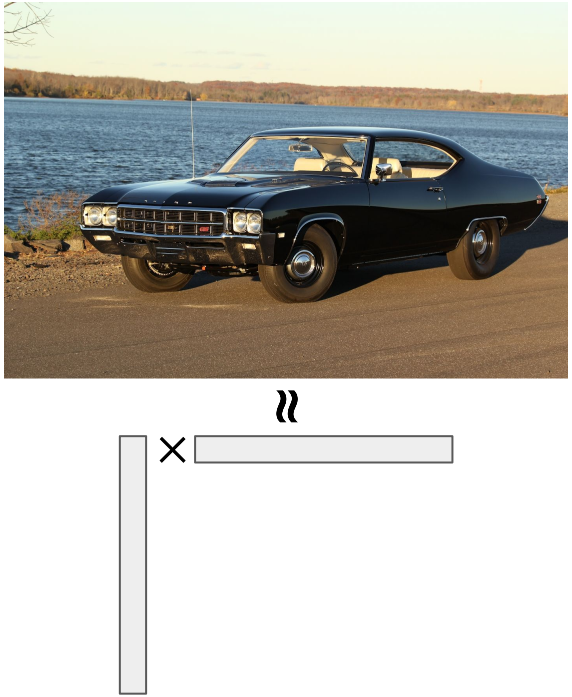
Figure 1: The overall experimental pipeline. EA: embedding alignment; Adv: adversarial training.

Hsu, Tzu-Ming Harry, et al. "Unsupervised multimodal representation learning across medical images and reports." *arXiv preprint arXiv:1811.08615* (2018).

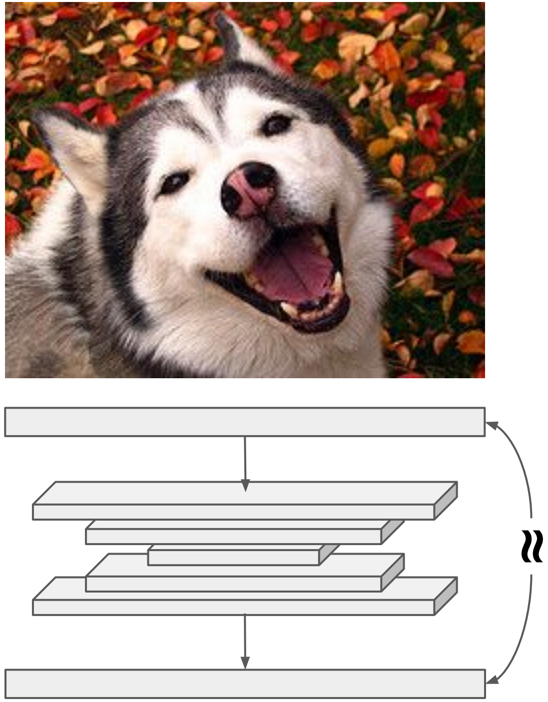
Representation: How can we learn?

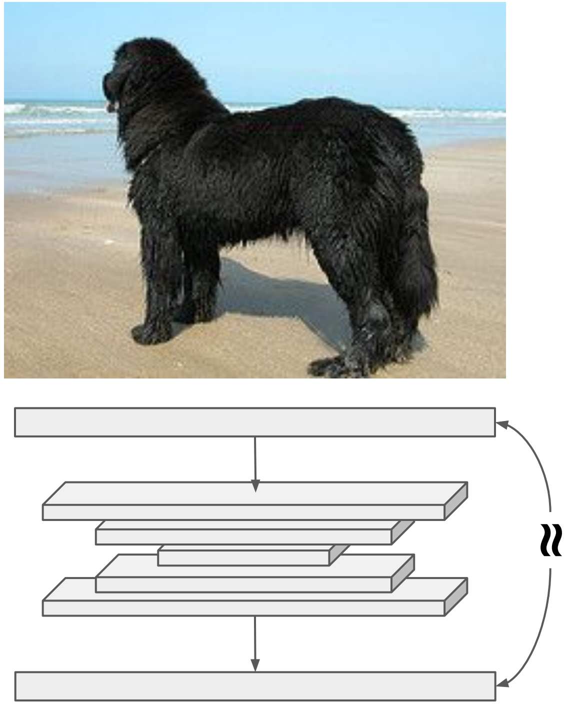


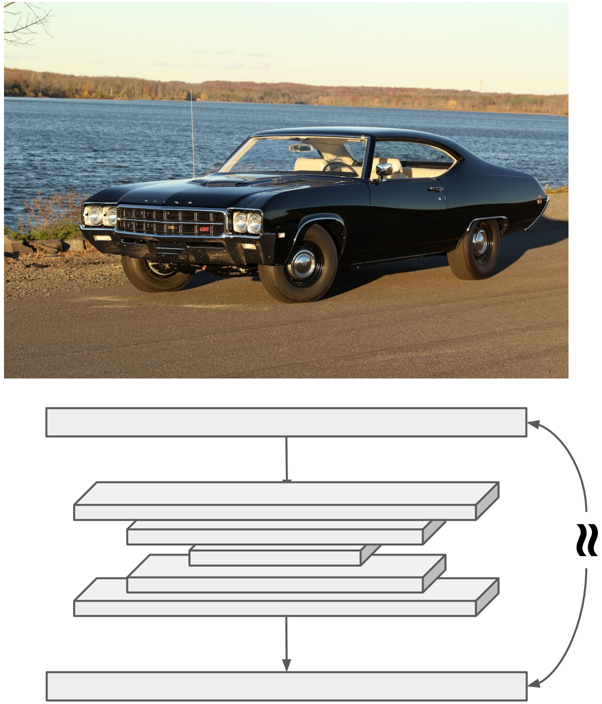




Representation: What can we do?







Representation: How can we learn?

“Dog, Husky, Grass, Leaves, Outside, Face, Teeth, Mottled, ...”

“Dog, Newfoundland, Sand, Waves, Outside, Beach, Shaggy, Black, ...”

“Car, 1969, Buick, GS 400, Outside, Road, Water, Hazy, Machine, Shiny ...”

Image size:
240 × 180

Find other sizes of this image:
All sizes - Small

Possible related search: ***newfoundland***

Welcome to Newfoundland and Labrador - Newfoundland and ...

https://www.newfoundlandlabrador.com/ ▾

Newfoundlandlabrador.com is your online resource for discovering everything you need to know about travelling to and around **Newfoundland** and Labrador.

Newfoundland and Labrador - Wikipedia

https://en.wikipedia.org/wiki/Newfoundland_and_Labrador ▾

Newfoundland and Labrador is the most easterly province of Canada. Situated in the country's Atlantic region, it comprises the island of **Newfoundland** and ...

Visually similar images

Report images

Pages that include matching images

I love Newfies!!! - Pinterest

Newfoundland dog

Dog breed

The Newfoundland dog is a large working dog. They can be either black, brown, or white-and-black. However, in the Dominion of Newfoundland, before it became part of the confederation of Canada, only black and Landseer coloured dogs were considered to be proper members of the breed. Wikipedia

Life span: 8 – 10 years

Weight: Male: 60–70 kg, Female: 45–55 kg

Height: Male: 69–74 cm, Female: 63–69 cm

Temperament: Sweet-Tempered, Trainable, Gentle

Colors: Black, Black & White, Grey, Brown

Origin: Canada, England

Newfoundland singers

View 25+ more

Dick Nolan

Ron Hynes

Harry Hibbs

Alan Doyle

Anita Best

People also search for

View 15+ more

St. Bernard

Bernese Mountain Dog

Leonberger

Great Dane

English Mastiff

Image size:
240 × 180

Find other sizes of this image:
All sizes - Small

Possible related search: ***newfoundland***

Welcome to Newfoundland and Labrador - Newfoundland and ...
https://www.newfoundlandlabrador.com/ ▾
Newfoundlandlabrador.com is your online resource for discovering everything you need to know about travelling to and around **Newfoundland** and Labrador.

Newfoundland and Labrador - Wikipedia
https://en.wikipedia.org/wiki/Newfoundland_and_Labrador ▾
Newfoundland and Labrador is the most easterly province of Canada. Situated in the country's Atlantic region, it comprises the island of **Newfoundland** and ...

Visually similar images

Report images

Pages that include matching images

I love Newfies!!! - Pinterest

Newfoundland dog

Dog breed

The Newfoundland dog is a large working dog. They can be either black, brown, or white-and-black. However, in the Dominion of Newfoundland, before it became part of the confederation of Canada, only black and Landseer coloured dogs were considered to be proper members of the breed. Wikipedia

Life span: 8 – 10 years

Weight: Male: 60–70 kg, Female: 45–55 kg

Height: Male: 69–74 cm, Female: 63–69 cm

Temperament: Sweet-Tempered, Trainable, Gentle

Colors: Black, Black & White, Grey, Brown

Origin: Canada, England

Newfoundland singers

View 25+ more

Dick Nolan | Ron Hynes | Harry Hibbs | Alan Doyle | Anita Best

People also search for

View 15+ more

St. Bernard | Bernese Mountain Dog | Leonberger | Great Dane | English Mastiff

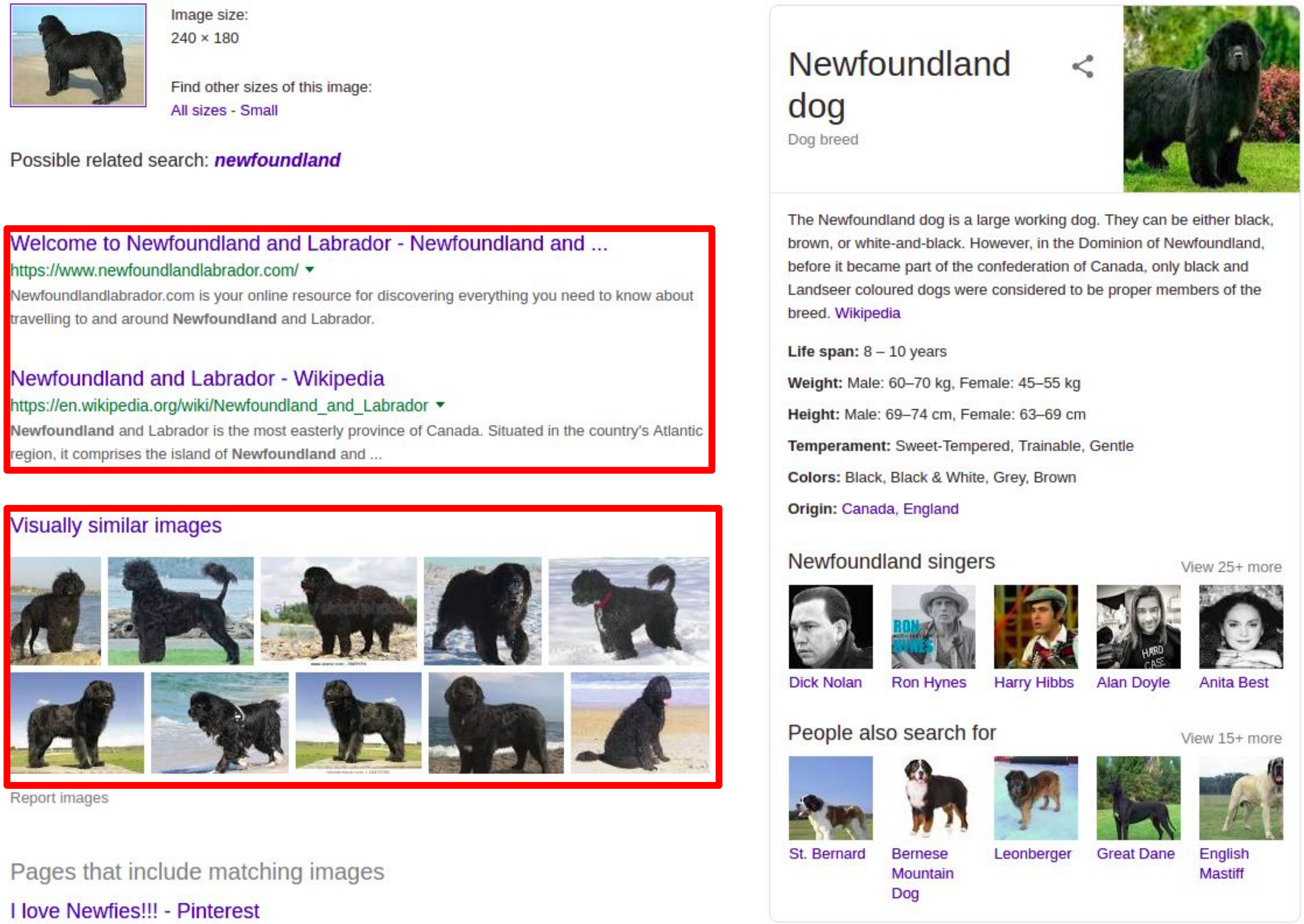


Image size:
240 × 180
Find other sizes of this image:
All sizes - Small
Possible related search: newfoundland
Welcome to Newfoundland and Labrador - Newfoundland and ...
https://www.newfoundlandlabrador.com/
Newfoundlandlabrador.com is your online resource for discovering everything you need to know about travelling to and around Newfoundland and Labrador.
Newfoundland and Labrador - Wikipedia
https://en.wikipedia.org/wiki/Newfoundland_and_Labrador
Newfoundland and Labrador is the most easterly province of Canada. Situated in the country's Atlantic region, it comprises the island of Newfoundland and ...
Visually similar images
Report images
Pages that include matching images
I love Newfies!!! - Pinterest
Newfoundland dog
Dog breed
The Newfoundland dog is a large working dog. They can be either black, brown, or white-and-black. However, in the Dominion of Newfoundland, before it became part of the confederation of Canada, only black and Landseer coloured dogs were considered to be proper members of the breed. Wikipedia
Life span: 8 – 10 years
Weight: Male: 60–70 kg, Female: 45–55 kg
Height: Male: 69–74 cm, Female: 63–69 cm
Temperament: Sweet-Tempered, Trainable, Gentle
Colors: Black, Black & White, Grey, Brown
Origin: Canada, England
Newfoundland singers
View 25+ more
Dick Nolan
Ron Hynes
Harry Hibbs
Alan Doyle
Anita Best
People also search for
View 15+ more
St. Bernard
Bernese Mountain Dog
Leonberger
Great Dane
English Mastiff

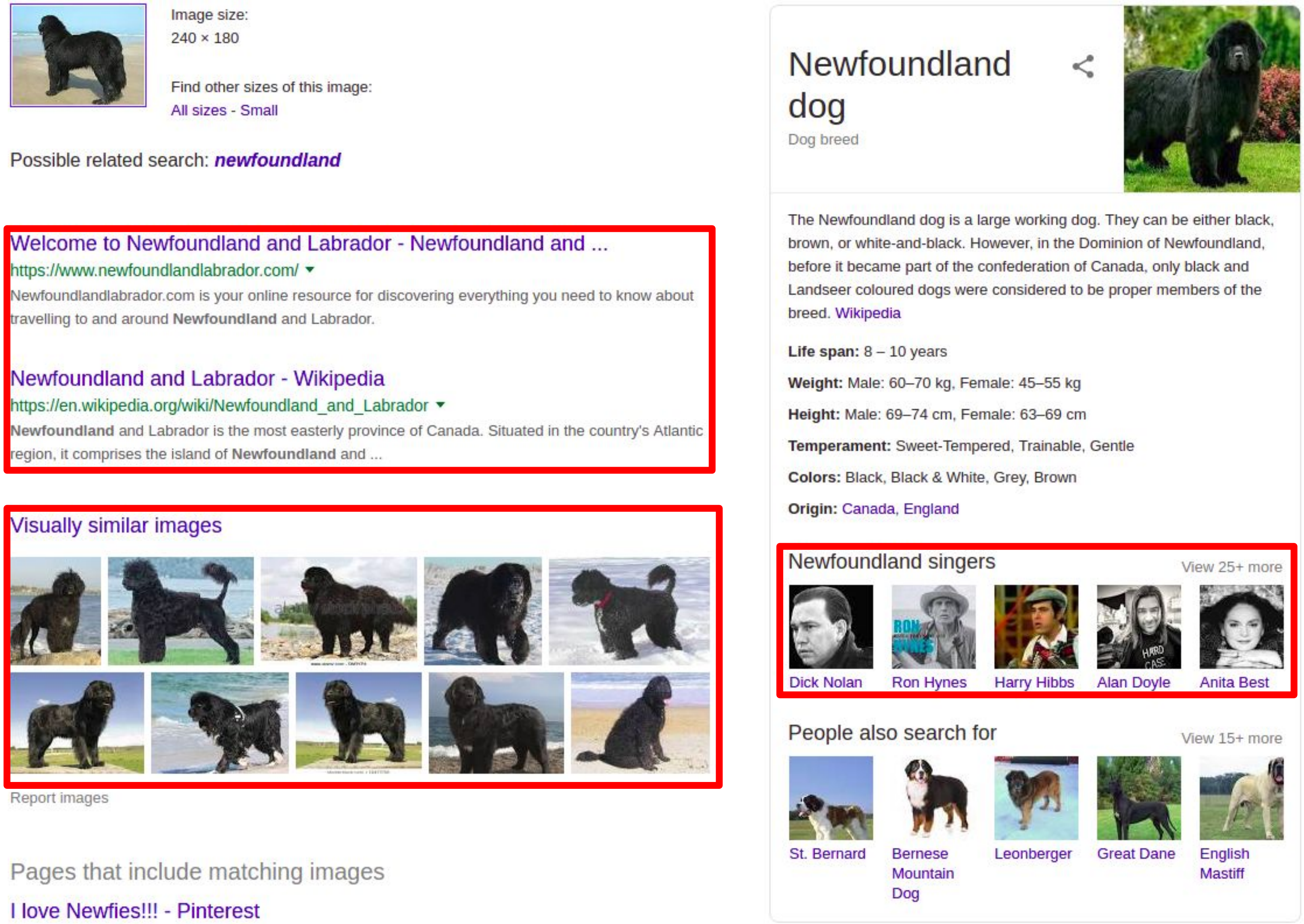


Image size:
240 × 180
Find other sizes of this image:
All sizes - Small
Possible related search: newfoundland
Welcome to Newfoundland and Labrador - Newfoundland and ...
https://www.newfoundlandlabrador.com/
Newfoundlandlabrador.com is your online resource for discovering everything you need to know about travelling to and around Newfoundland and Labrador.
Newfoundland and Labrador - Wikipedia
https://en.wikipedia.org/wiki/Newfoundland_and_Labrador
Newfoundland and Labrador is the most easterly province of Canada. Situated in the country's Atlantic region, it comprises the island of Newfoundland and ...
Visually similar images
Report images
Pages that include matching images
I love Newfies!!! - Pinterest
Newfoundland dog
Dog breed
The Newfoundland dog is a large working dog. They can be either black, brown, or white-and-black. However, in the Dominion of Newfoundland, before it became part of the confederation of Canada, only black and Landseer coloured dogs were considered to be proper members of the breed. Wikipedia
Life span: 8 – 10 years
Weight: Male: 60–70 kg, Female: 45–55 kg
Height: Male: 69–74 cm, Female: 63–69 cm
Temperament: Sweet-Tempered, Trainable, Gentle
Colors: Black, Black & White, Grey, Brown
Origin: Canada, England
Newfoundland singers
View 25+ more
Dick Nolan
Ron Hynes
Harry Hibbs
Alan Doyle
Anita Best
People also search for
View 15+ more
St. Bernard
Bernese Mountain Dog
Leonberger
Great Dane
English Mastiff

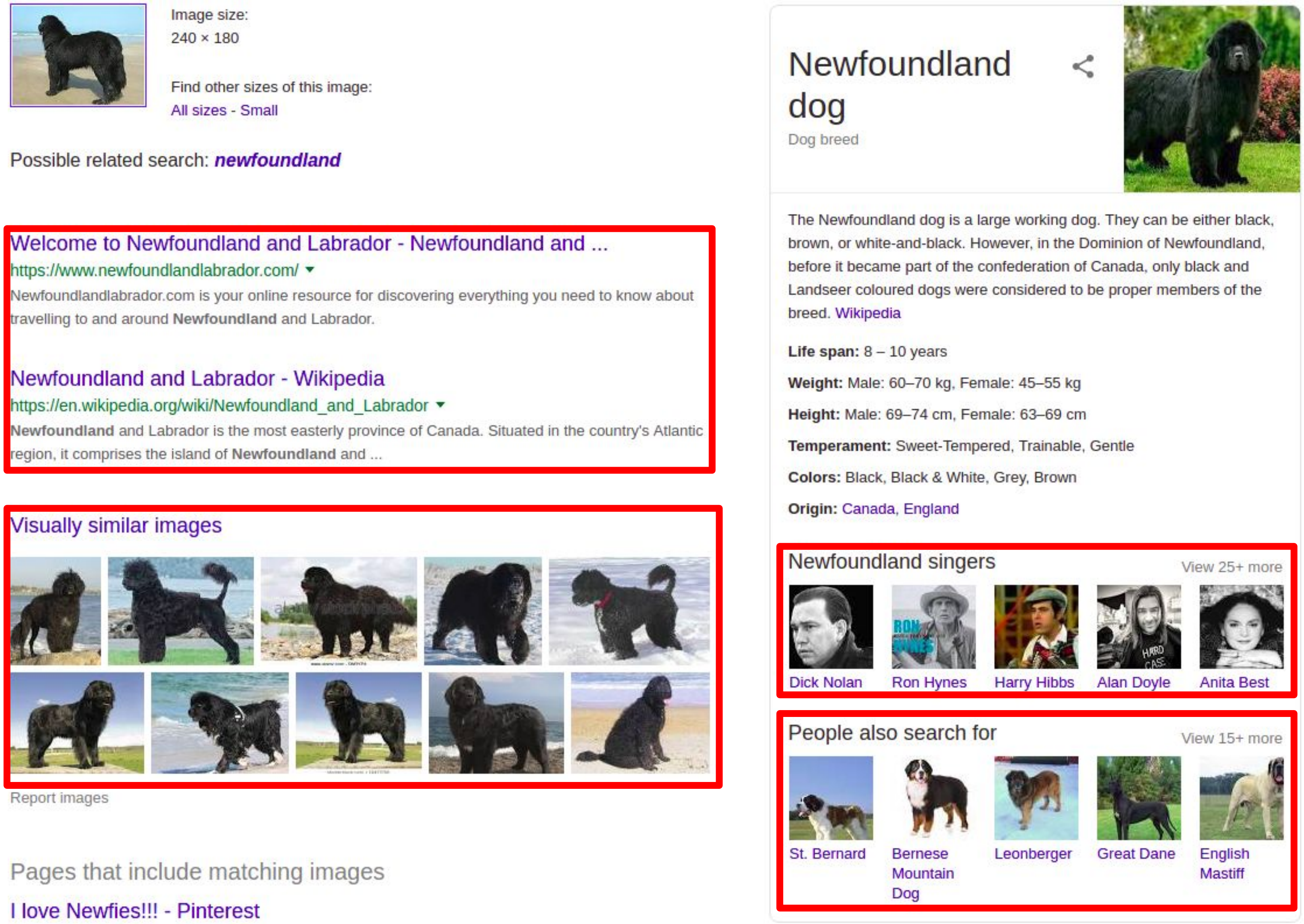


Image size:
240 × 180
Find other sizes of this image:
All sizes - Small
Possible related search: newfoundland
Welcome to Newfoundland and Labrador - Newfoundland and ...
https://www.newfoundlandlabrador.com/
Newfoundlandlabrador.com is your online resource for discovering everything you need to know about travelling to and around Newfoundland and Labrador.
Newfoundland and Labrador - Wikipedia
https://en.wikipedia.org/wiki/Newfoundland_and_Labrador
Newfoundland and Labrador is the most easterly province of Canada. Situated in the country's Atlantic region, it comprises the island of Newfoundland and ...
Visually similar images
Report images
Pages that include matching images
I love Newfies!!! - Pinterest
Newfoundland dog
Dog breed
The Newfoundland dog is a large working dog. They can be either black, brown, or white-and-black. However, in the Dominion of Newfoundland, before it became part of the confederation of Canada, only black and Landseer coloured dogs were considered to be proper members of the breed. Wikipedia
Life span: 8 – 10 years
Weight: Male: 60–70 kg, Female: 45–55 kg
Height: Male: 69–74 cm, Female: 63–69 cm
Temperament: Sweet-Tempered, Trainable, Gentle
Colors: Black, Black & White, Grey, Brown
Origin: Canada, England
Newfoundland singers
View 25+ more
Dick Nolan
Ron Hynes
Harry Hibbs
Alan Doyle
Anita Best
People also search for
View 15+ more
St. Bernard
Bernese Mountain Dog
Leonberger
Great Dane
English Mastiff

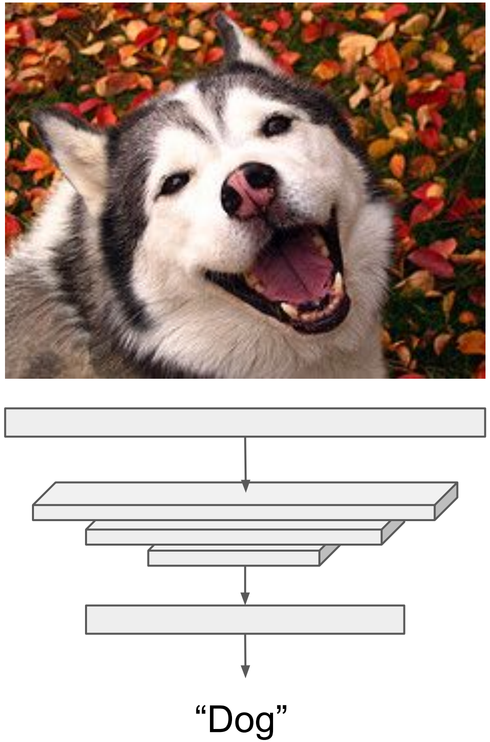
Representation: How can we learn?

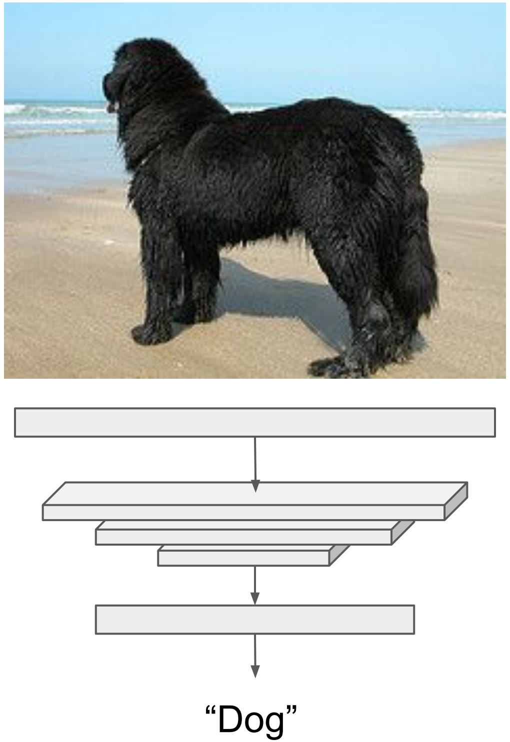
“Dog, Husky, Grass, Leaves, Outside, Face, Teeth, Mottled, ...”

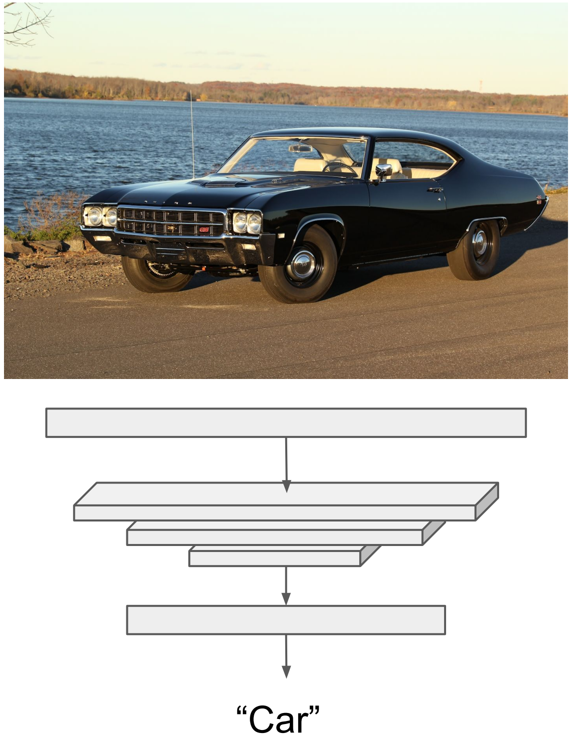
“Dog, Newfoundland, Sand, Waves, Outside, Beach, Shaggy, Black, ...”

“Car, 1969, Buick, GS 400, Outside, Road, Water, Hazy, Machine, Shiny ...”

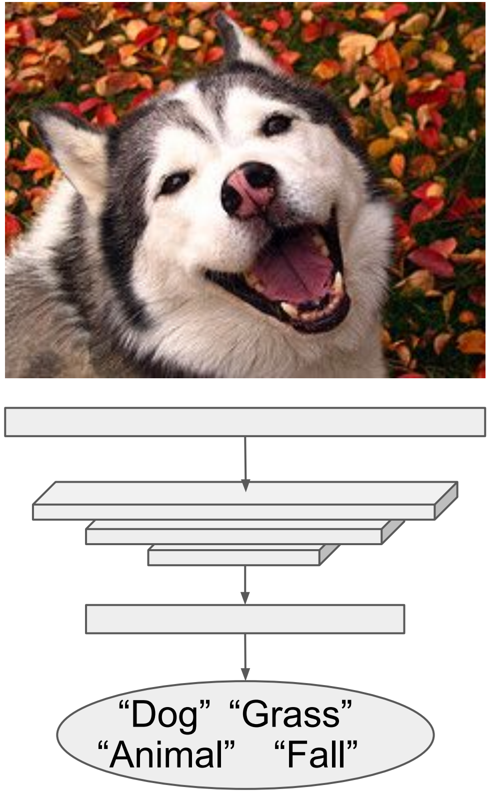
Representation: What can we do?

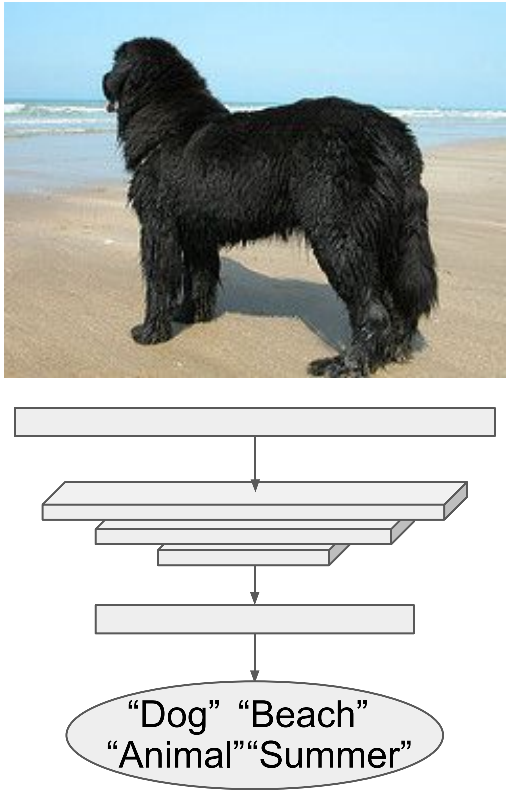


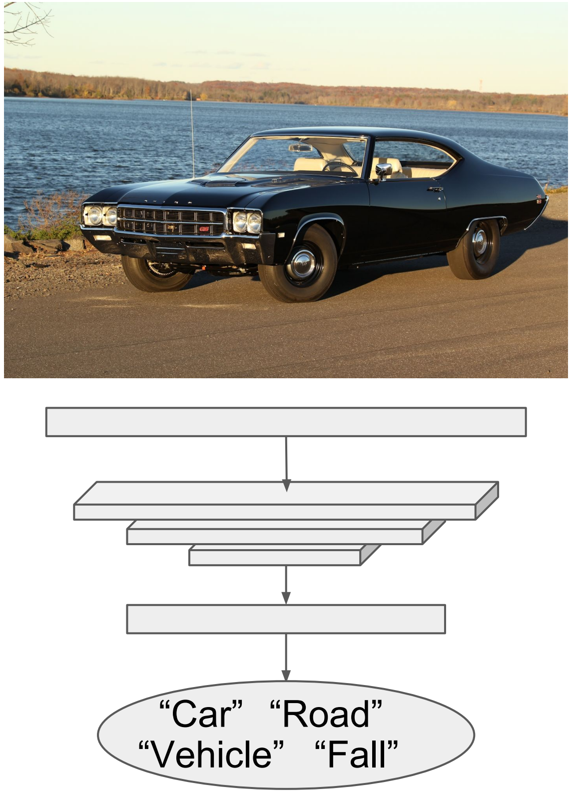




Representation: What can we do?







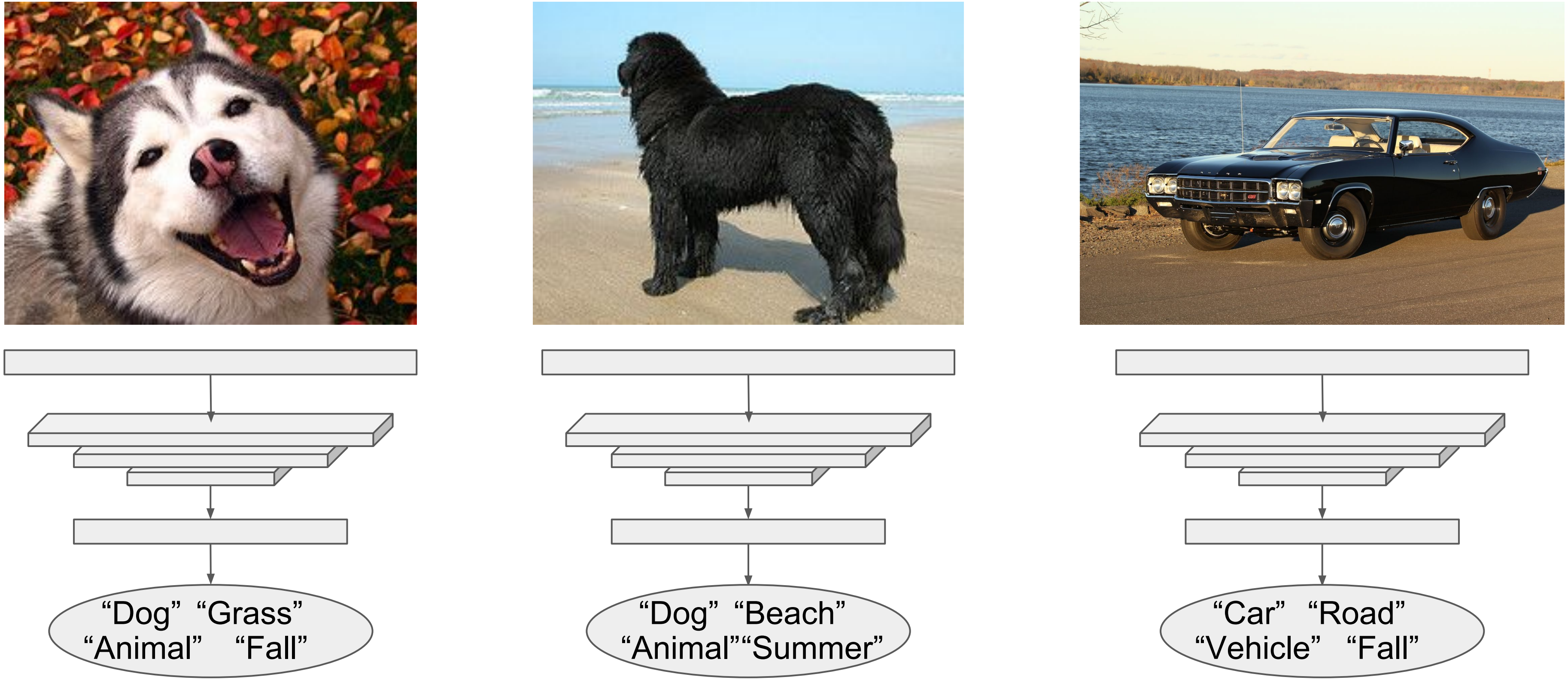
Representation: How can we learn?

“Dog, Husky, Grass, Leaves, Outside, Face, Teeth, Mottled, ...”

“Dog, Newfoundland, Sand, Waves, Outside, Beach, Shaggy, Black, ...”

“Car, 1969, Buick, GS 400, Outside, Road, Water, Hazy, Machine, Shiny ...”

Representation: What can we do?



DeepCluster: Why bother with labels?

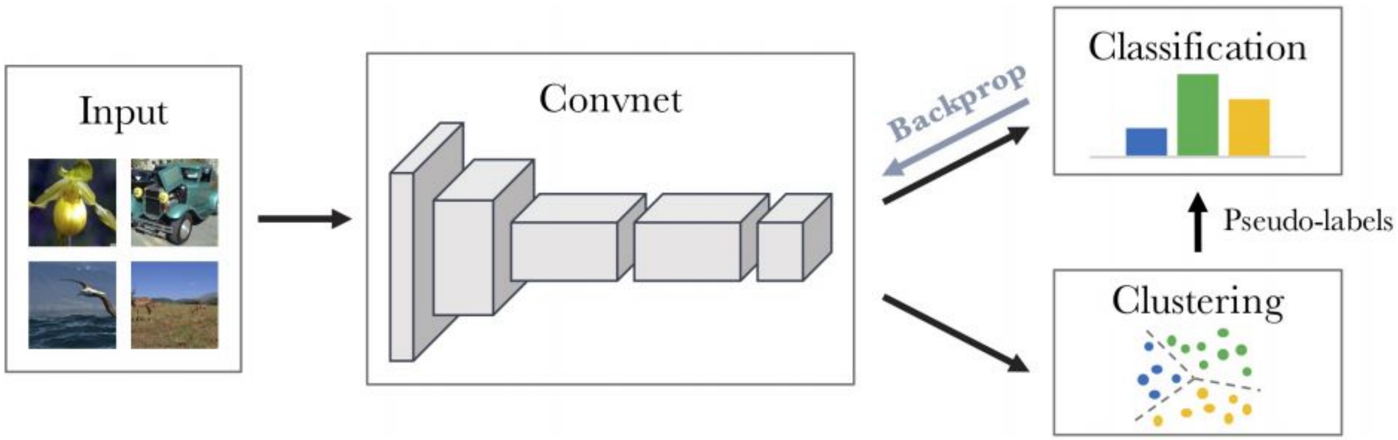


Fig. 1: Illustration of the proposed method: we iteratively cluster deep features and use the cluster assignments as pseudo-labels to learn the parameters of the convnet.

Caron, Mathilde, et al. "Deep clustering for unsupervised learning of visual features." Proceedings of the European Conference on Computer Vision (ECCV). 2018.

Representation Learning in Action: Multitask Learning

Multi-task Prediction of Disease Onsets from Longitudinal Lab Tests

Narges Razavian, Jake Marcus, David Sontag
Courant Instit
{

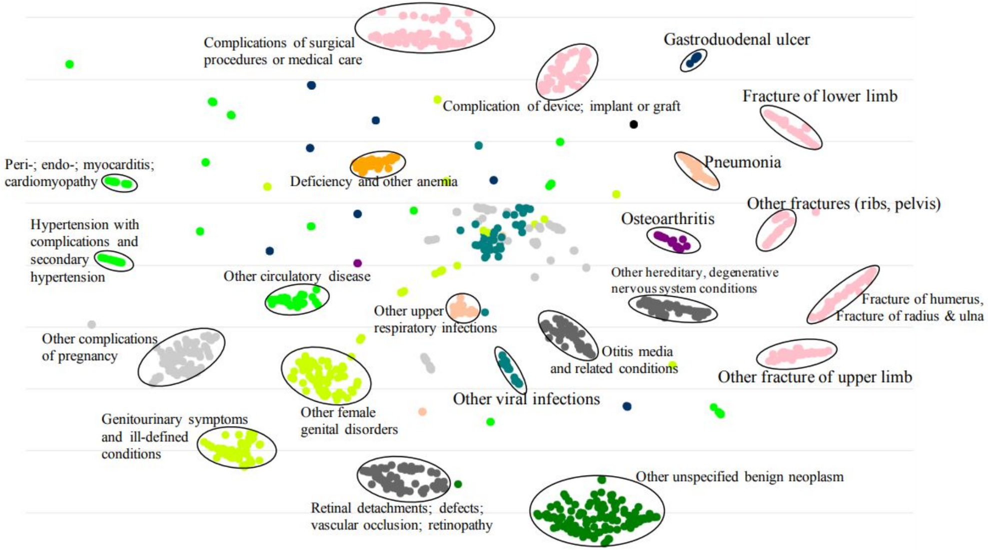
Multitask Learning and Benchmarking with Clinical Time Series Data

Hrayr Harutyunyan[1], Hrant Khachatrian[2,3], David C. Kale[1], Greg Ver Steeg[1], and Aram Galstyan[1]
nited States of America

MoleculeNet: a benchmark for molecular machine learning†

Zhenqin Wu, ‡[a] Bharath Ramsundar, ‡[b] Evan N. Feinberg, §[c] Joseph Gomes, §[a] Caleb Geniesse,[c] Aneesh S. Pappu,[b] Karl Leswing[d] and Vijay Pande*[a]

Representation Learning in Action: Clustering



(a) Scatterplot of the final representations $\mathbf{g}_i$'s of GRAM+

Choi, Edward, et al. "GRAM: graph-based attention model for healthcare representation learning." *Proceedings of the 23rd ACM SIGKDD International Conference on Knowledge Discovery and Data Mining*. ACM, 2017.

Representation Learning in Action: Clustering

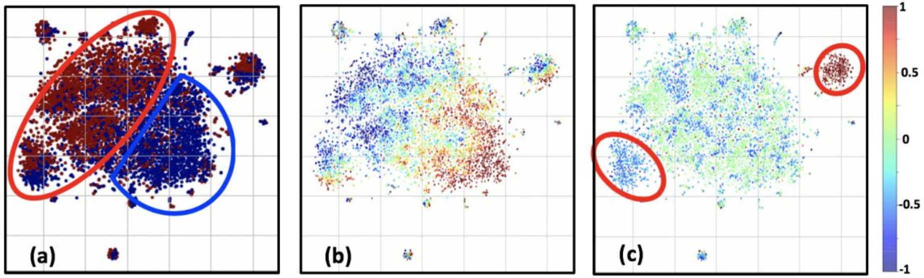


Figure 3: tSNE on context vectors of test dataset from BSS model colored by (a) red: positive examples and blue: negative examples, (b) average systemic diastolic blood pressure; and (c) average central venous pressure.

Dhamala, Jwala, et al. "Multivariate Time-Series Similarity Assessment via Unsupervised Representation Learning and Stratified Locality Sensitive Hashing: Application to Early Acute Hypotensive Episode Detection." *IEEE Sensors Letters* 3.1 (2019): 1-4.

Representation Learning in Action: Anomaly Detection

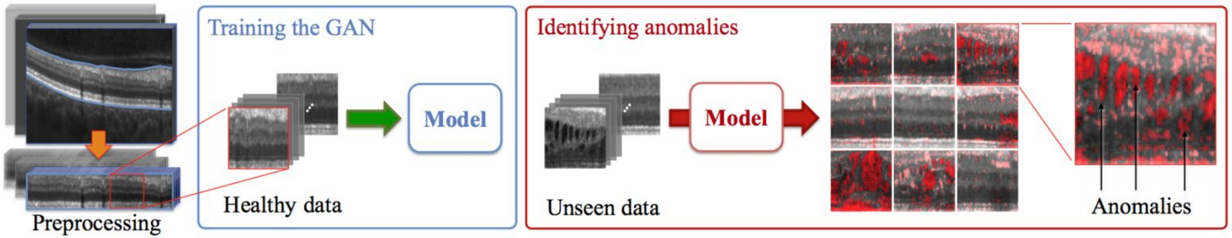


Fig. 1. Anomaly detection framework. The preprocessing step includes extraction and flattening of the retinal area, patch extraction and intensity normalization. Generative adversarial training is performed on healthy data and testing is performed on both, unseen healthy cases and anomalous data.

Schlegl, Thomas, et al. "Unsupervised anomaly detection with generative adversarial networks to guide marker discovery." *International Conference on Information Processing in Medical Imaging*. Springer, Cham, 2017.

Representation Learning in Action: Anomaly Detection

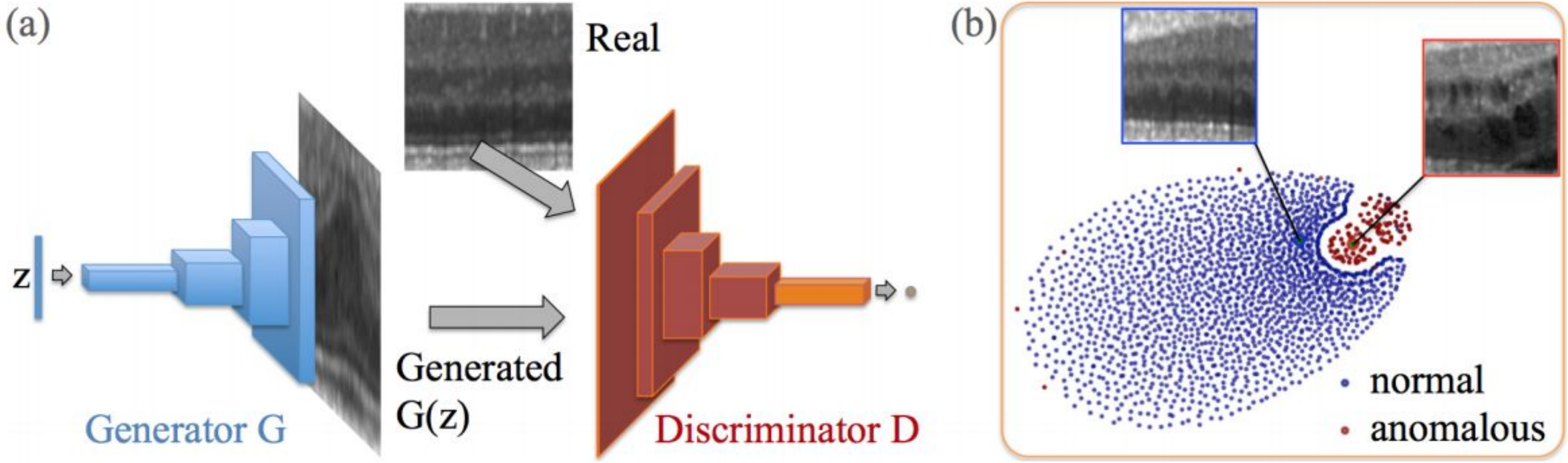


Fig. 2. (a) Deep convolutional generative adversarial network. (b) t-SNE embedding of normal (blue) and anomalous (red) images on the feature representation of the last convolution layer (orange in (a)) of the discriminator.

Key Points for Healthcare

- Representations can normalize.
- Generalization to unseen tasks is critical (e.g., patient subtyping).
- Representations can aid in interpretability.
- Representations can span many modalities.

What can you do with a representation?

1. Confounder Adjustment: https://arxiv.org/pdf/1811.06498.pdf

Evaluating a Representation

How can we ensure our representations are generalizable?

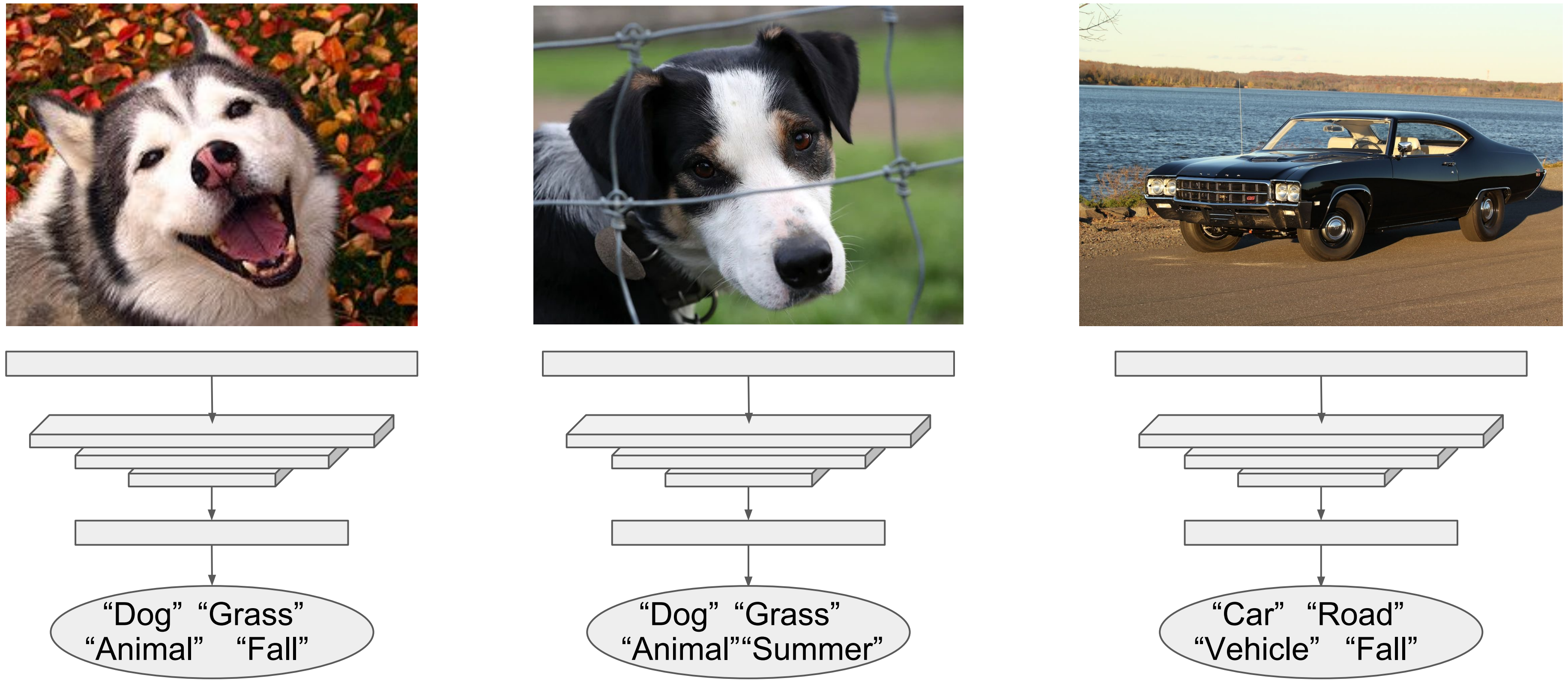
Evaluating a Representation

How can we ensure our representations are **generalizable**?

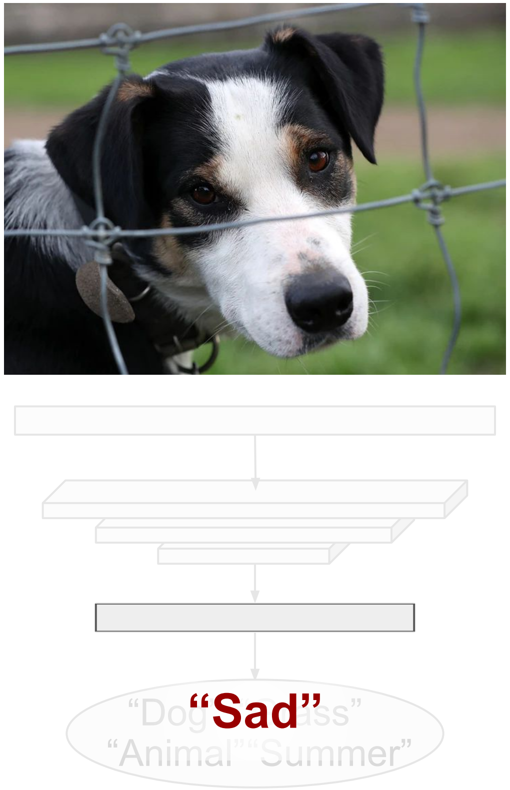
To new data

To new problems

Generalizable Representations

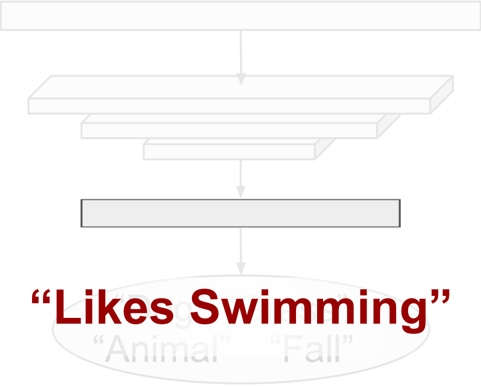


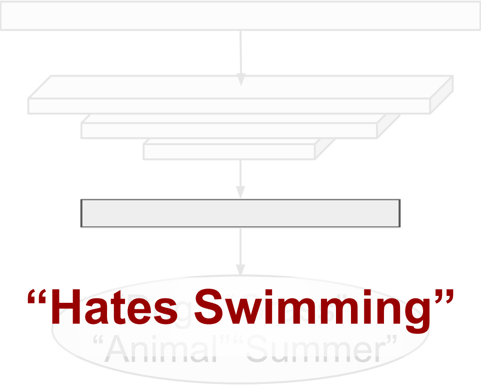
Generalizable Representations?

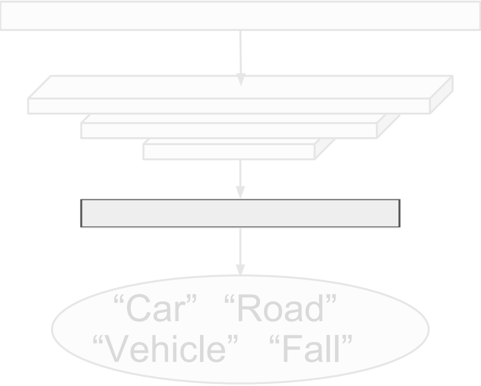


Fully Generalizable Representations are Not Possible

Data may not be enough

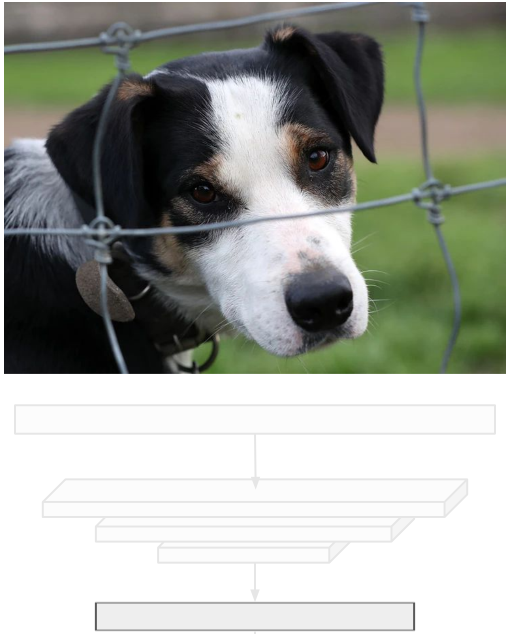






Data may not be enough

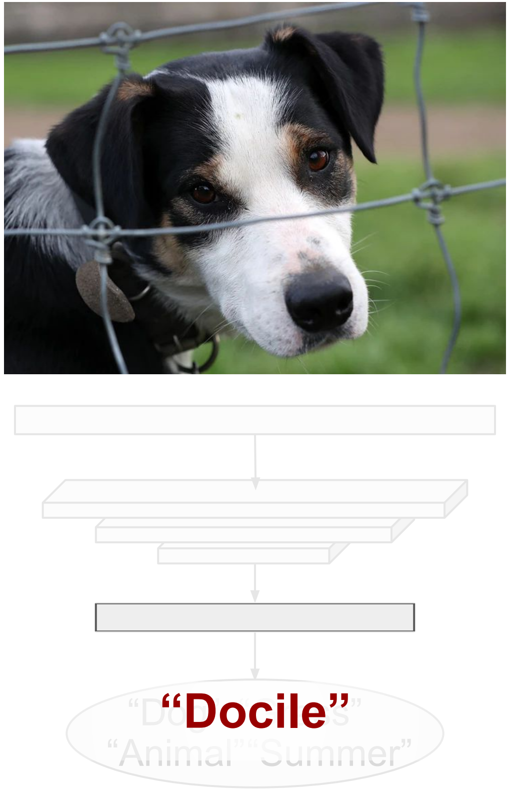
“Likes Swimming”



“Hates Swimming”

“Why did you try to take your car swimming???”

Task may be totally out of the box



We can still do a lot

True tests of evaluation/Parting thoughts

- Transfer Learning

True tests of evaluation/Parting thoughts

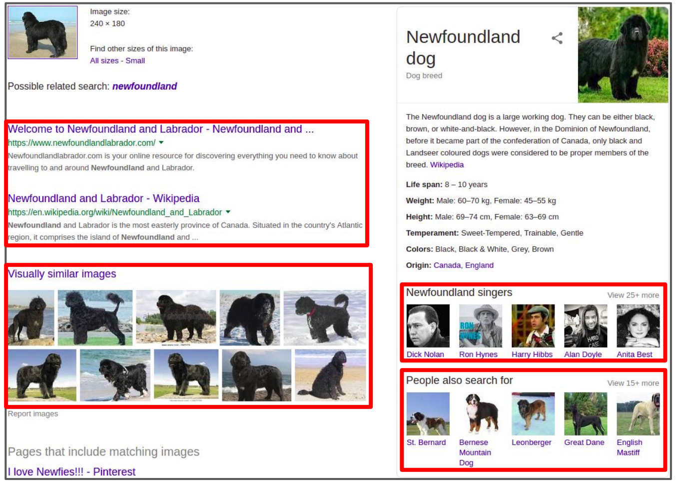
- Transfer Learning
- Operationalization (leave-one-task-out) possible, but questions remain:
  - Test under reduced dataset sizes
  - “Current tasks” is not a random sample of “possible tasks”

True tests of evaluation/Parting thoughts

- Transfer Learning
- Operationalization (leave-one-task-out) possible, but questions remain:
  - Test under reduced dataset sizes
  - “Current tasks” is not a random sample of “possible tasks”
- How do we evaluate multi-dimensional notions of similarity?

True tests of evaluation/Parting thoughts

- Transfer Learning
- Operationalization (leave-one-task-out) possible, but questions remain:
  - Test under reduced dataset sizes
  - “Current tasks” is not a random sample of “possible tasks”
- How do we evaluate multi-dimensional notions of similarity?



Thank you for your attention

- Links to all papers mentioned available with slides
- Get in touch: mmd@mit.edu
- Questions?

Representation Learning in the Literature

Representation Learning: A Review and New Perspectives

Yoshua Bengio[†], Aaron Courville, and Pascal Vincent[†]
Department of computer science and operations research, U. Montreal
† also, Canadian Institute for Advanced Research (CIFAR)

Why do we care?

- Missing data is inconvenient.
- Missing data is not going away.
- **Missing data is informative (or confounding).**