Medicine powered by AI

06/04/19

Maciej Pajak
maciej.pajak@eu.medical.canon
AI Scientist

AI Research Team
Canon Medical Research Europe

Canon Medical Research Europe

1994
Company Founded by two University of Edinburgh Graduates

1999
Started working with Toshiba

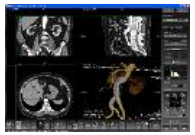
2001
1st Ultrasound product for Toshiba

2003
1st Technology Framework R&D for Toshiba

2009
Acquired by Toshiba

2011
R&D Grant from Scottish Govt

2018
Rebranded as a Canon Medical Research



Global R&D Organisation

Canon Medical Systems Corporation

Edinburgh

Marseille

Dailan

Beijing

Nasu & Yokohama

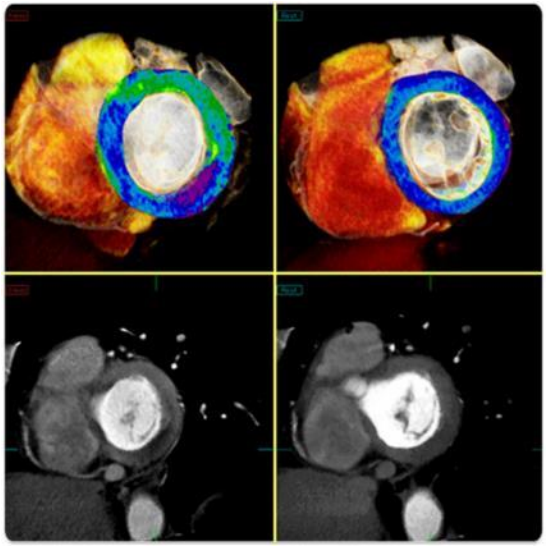
Minneapolis

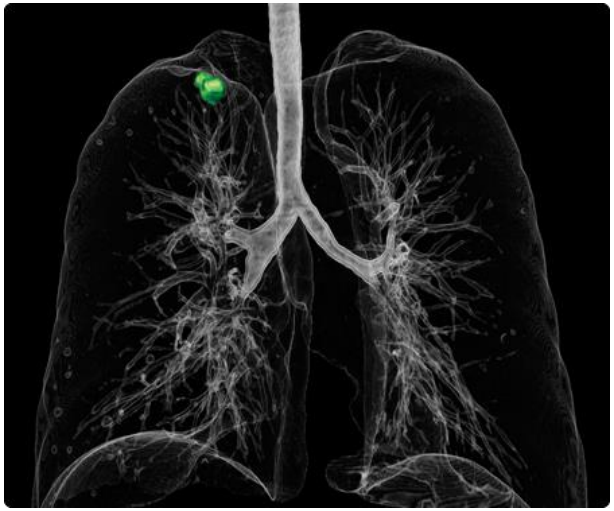
Chicago

Waterloo

R&D Collaboration in Action

Visualisation





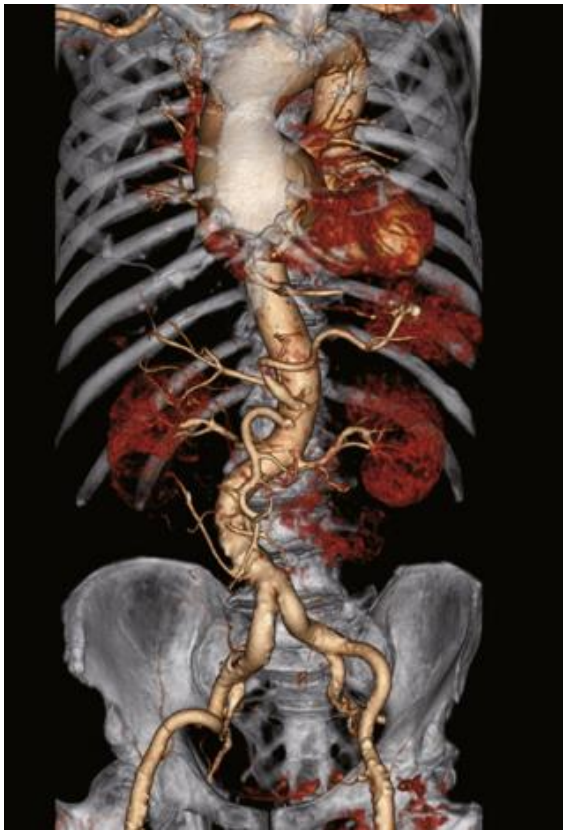
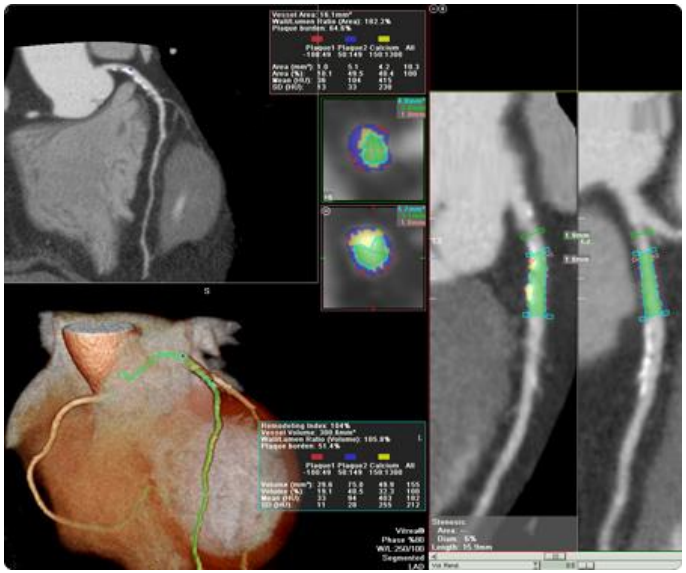
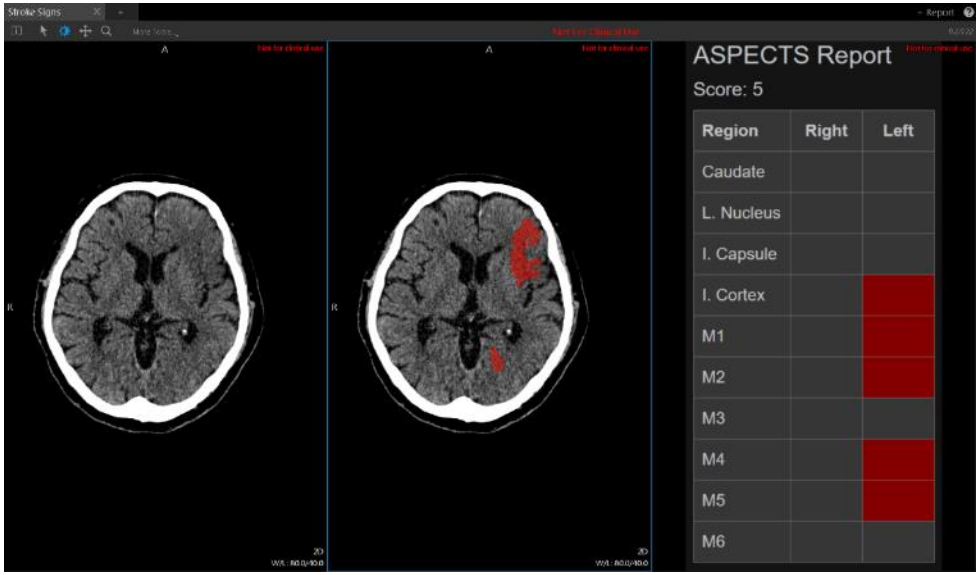
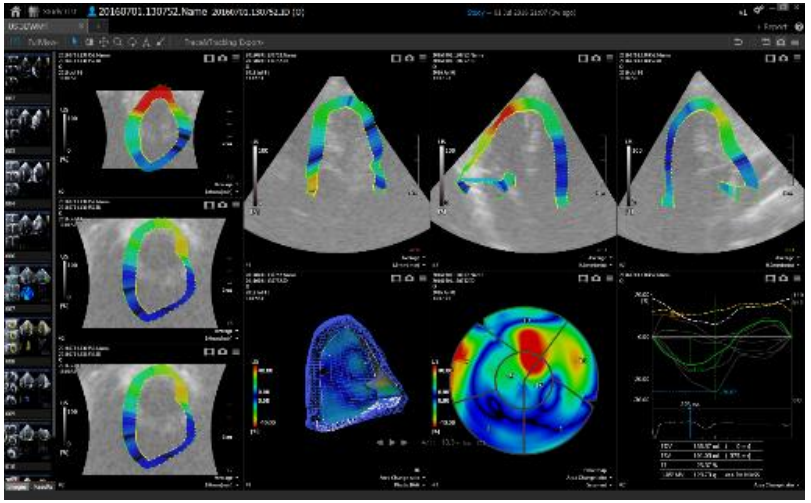


Image Analysis



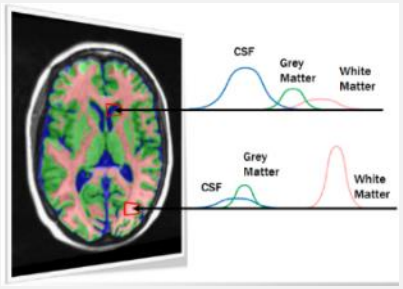




AI at Canon

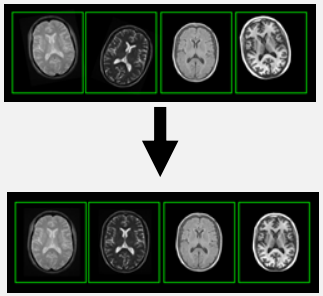
2009: GM/WM/CSF tissue classification in MR Brain

Gaussian mixture model



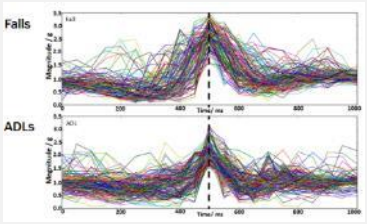
2015: Image registration

Fully convolutional NN



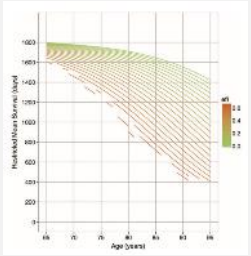
2015: Fall detection from accelerometer

1D Neural Network



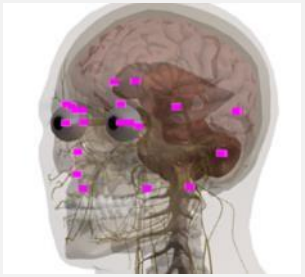
2018: Frailty measurement from coded medical data

Gradient boosting classifier



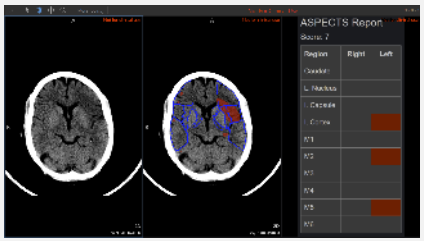
2012: Anatomical landmark detection

Random forest (product), Fully convolutional NN (R&D)



2016: Stroke signs detection

Fully convolutional NN



2017: Rule-based information retrieval

Rule-based with ontologies

2019: Natural language inference for healthcare text

Transformer NN

Canon

iCAIRD

The Industrial Centre for Artificial Intelligence Research in Digital Diagnostics

Alison Murray
Professor of Radiology, Director of SINAPSE, University of Aberdeen, NHS Grampian

Peter Hamilton
Leader Image Analytics, Philips Digital Pathology Solutions, Hon Professor of Tissue Imaging, QUB, Belfast

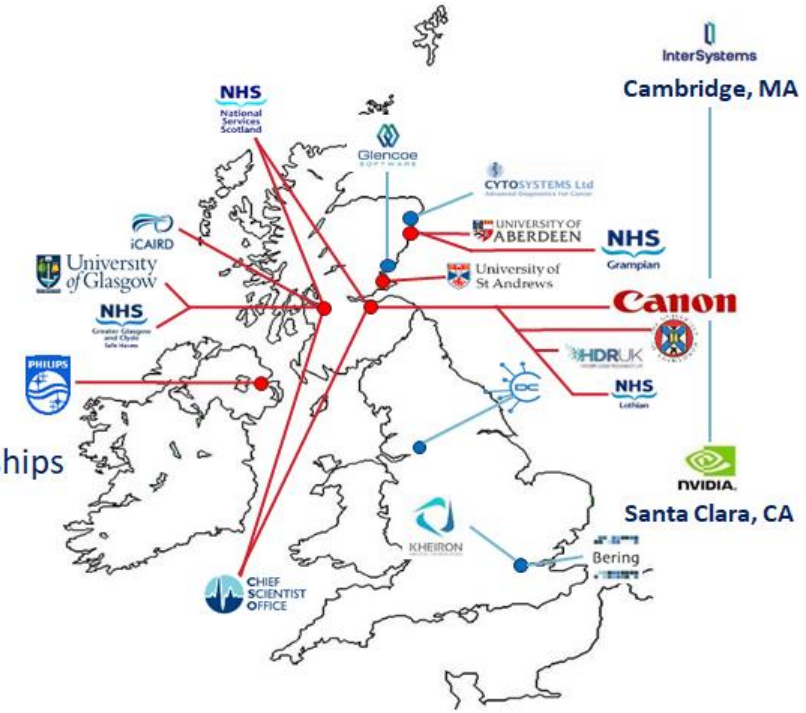
Andy Smout
Vice President Research, Canon Medical, Edinburgh

Colin McCowan
Professor of Health Informatics, University of Glasgow and Glasgow Safe Haven

David Harrison
Professor of Pathology, Universities of St Andrews & Edinburgh, NHS Lothian

Key features

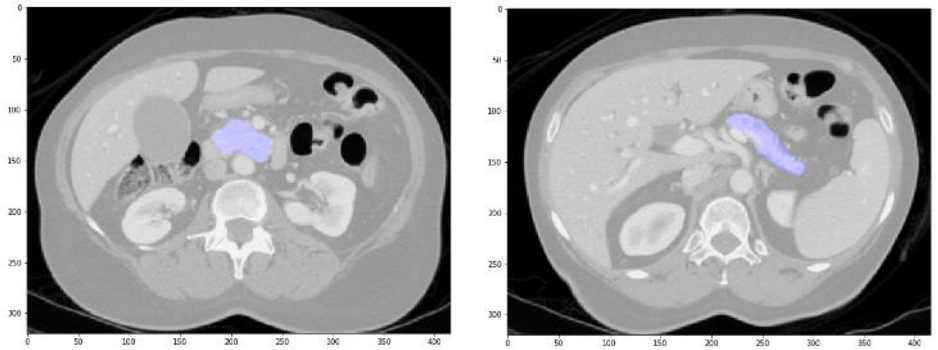
- Tried & tested partnerships
- NHS NSS Board
- Safe havens & HDRUK
- SINAPSE
- National PACS
- Direct link to clinicians



Read more

The biggest challenge in medical AI

[Wikimedia]



[Medical Decathlon]

#{Pictures of cats}>>#{Labelled medical scans}

Google 'betrays patient trust' with DeepMind Health move

Moving healthcare subsidiary into main company breaks pledge that 'data will not be connected to Google accounts'

[Guardian]

SCIENCE

IT'S EASIER TO DONATE YOUR BODY TO SCIENCE THAN YOUR MEDICAL RECORDS

Handing over your digital health data is still nearly impossible

[The Verge]

Jury of citizens ponder tricky question about health records

To what extent should patients control access to patient records?

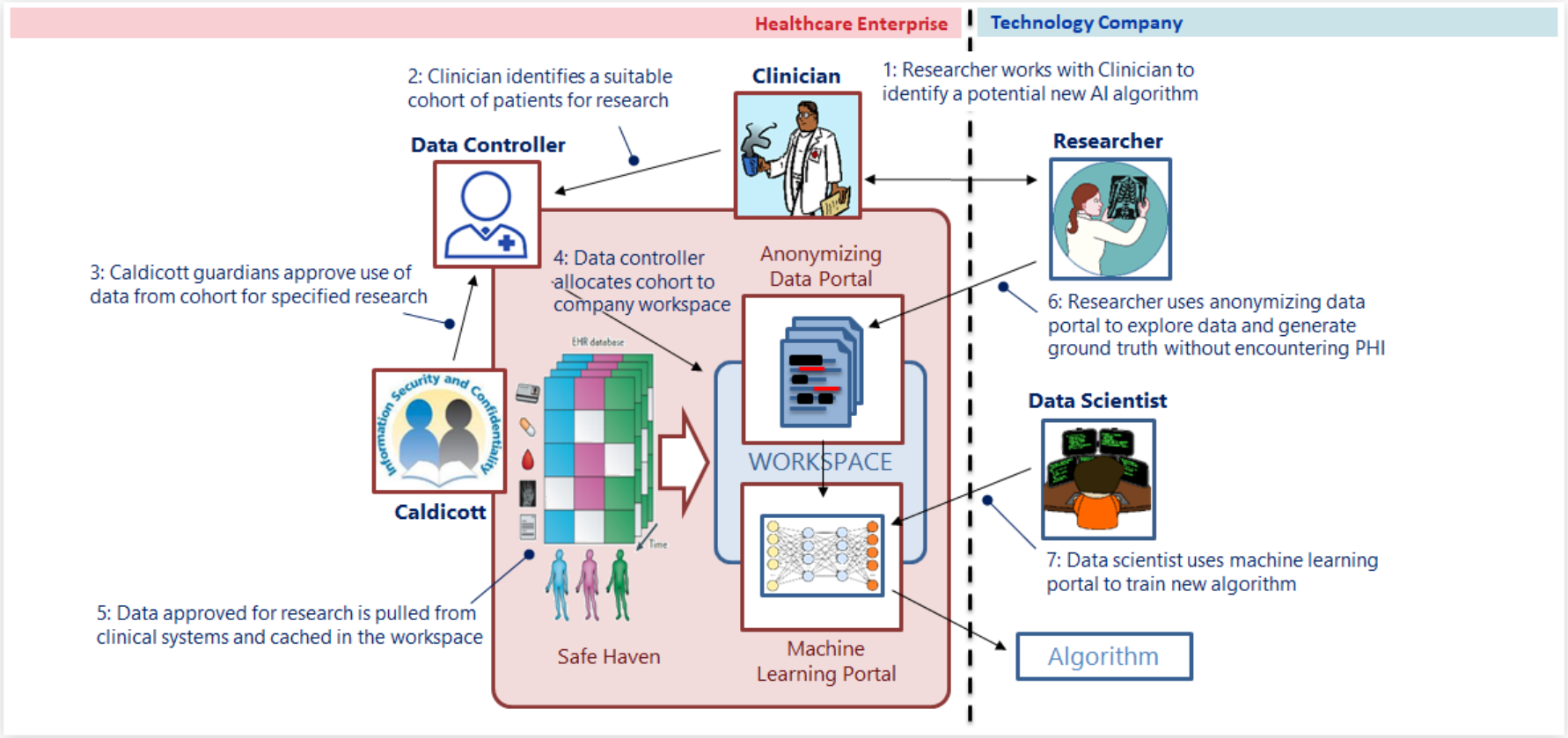
It's a difficult question that is at the heart of the public controversy which arose in 2014 around care data and the "selling" of hospital records to private companies. Researchers, epidemiologists, managers and many others rely on getting access to detailed information in patient records to improve the effectiveness and safety of future health services. But many people feel uneasy about others delving into their patient record, even if it is for a good cause.

[Health e-Research Centre]

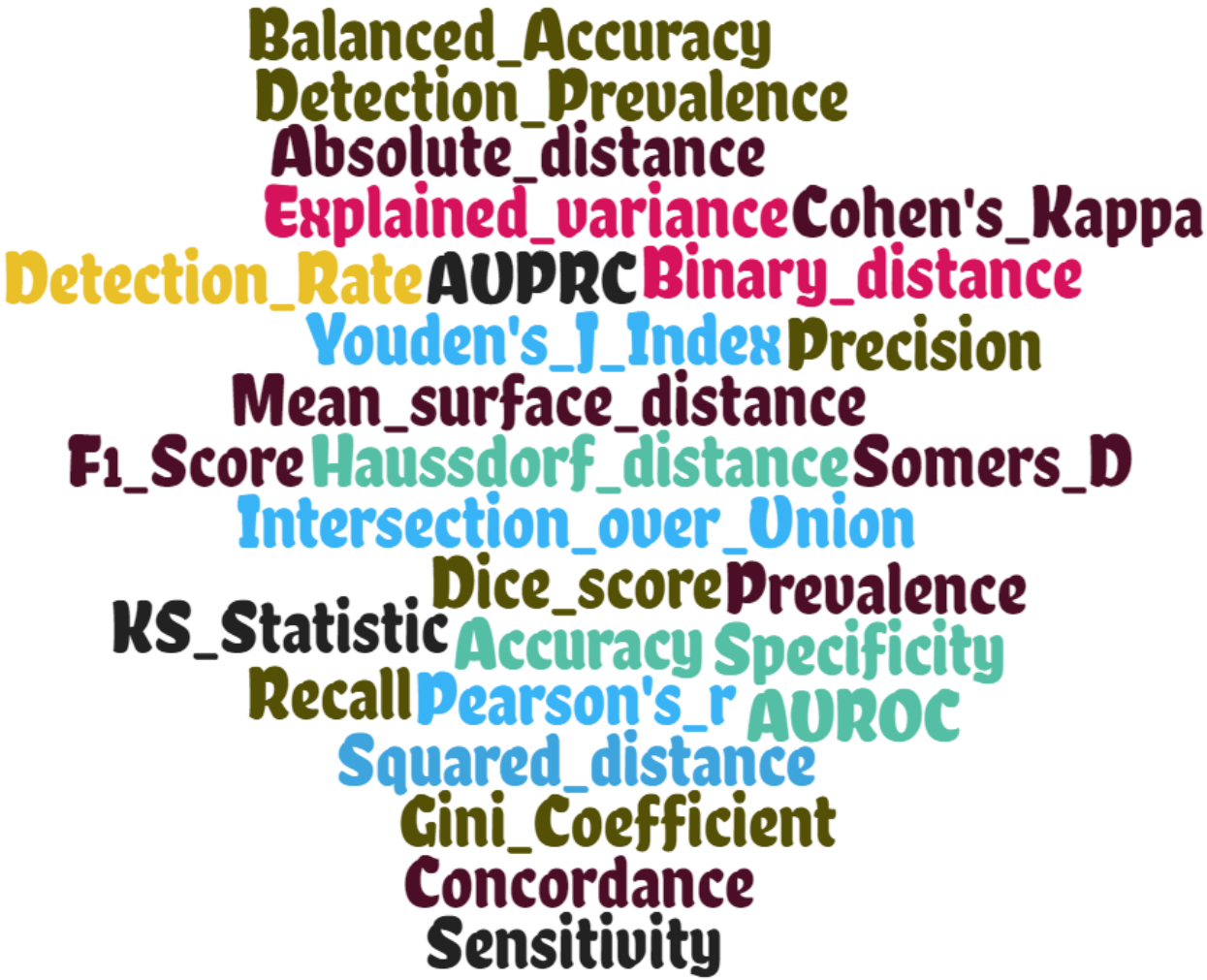
Canon

SHAIP

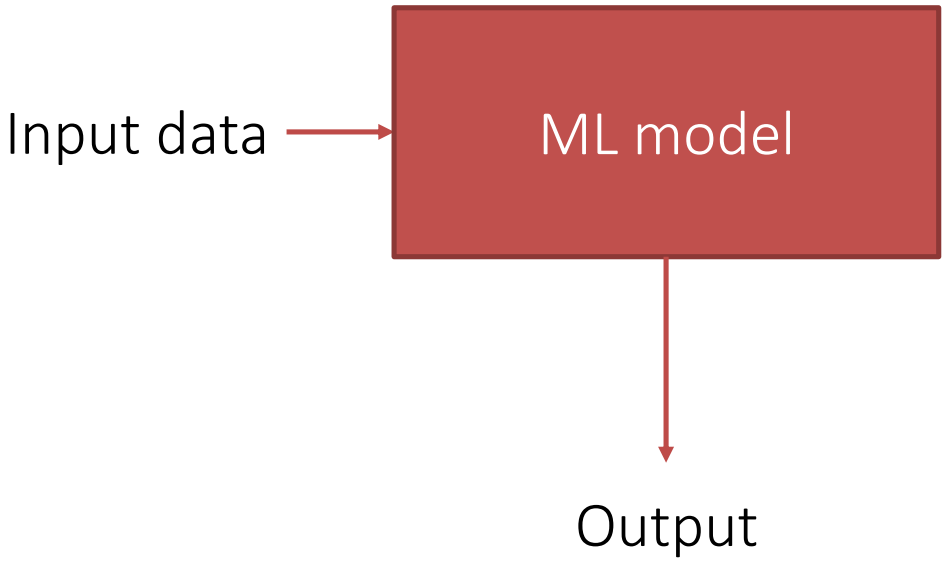
Save Haven AI Platform



The ~~science~~ art of measuring performance

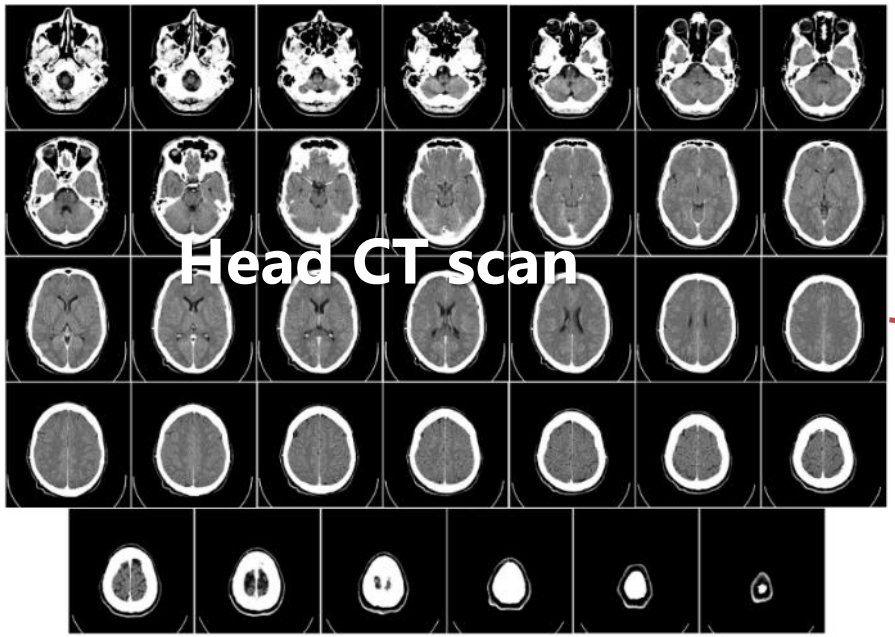


Note



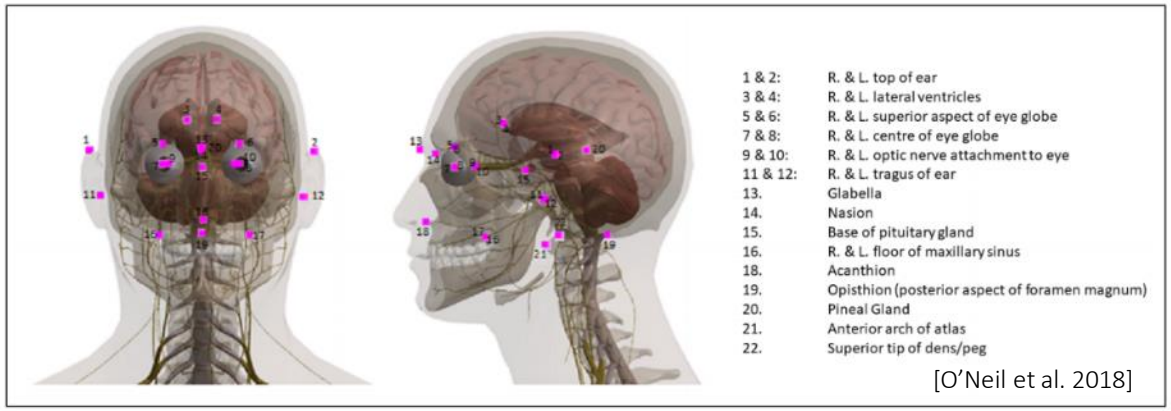
How good is good enough?

- Prediction of landmark (x,y,z) coordinates



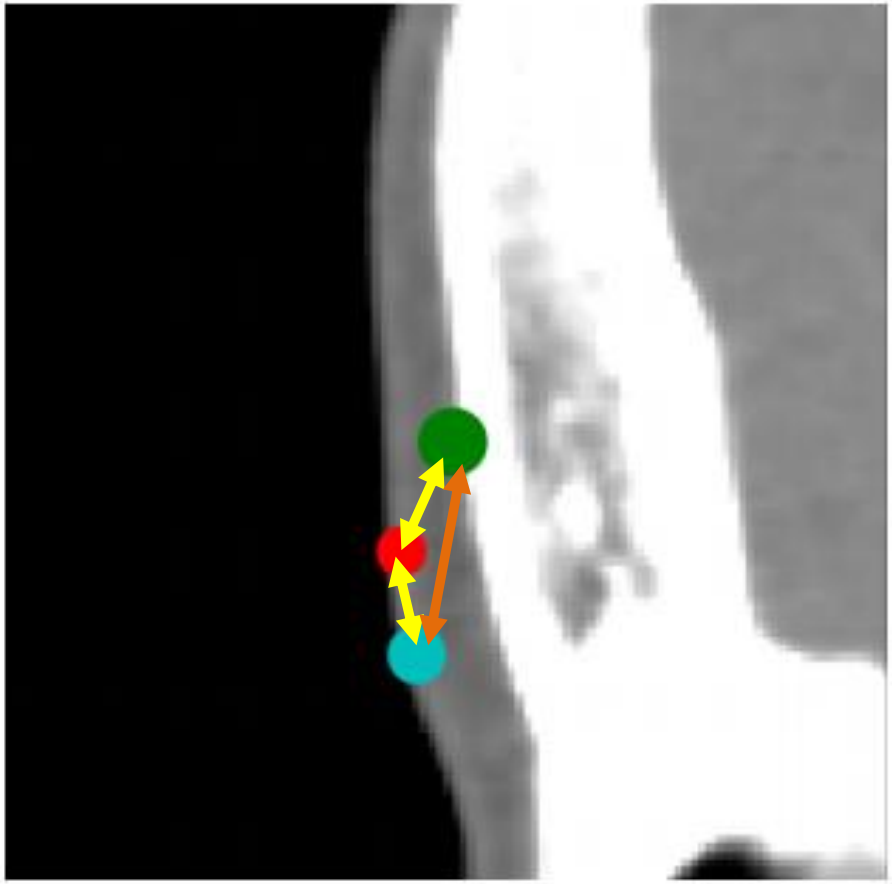
ML model

[Radiology Dept, Uppsala University Hospital]



[O'Neil et al. 2018]

How good is good enough?

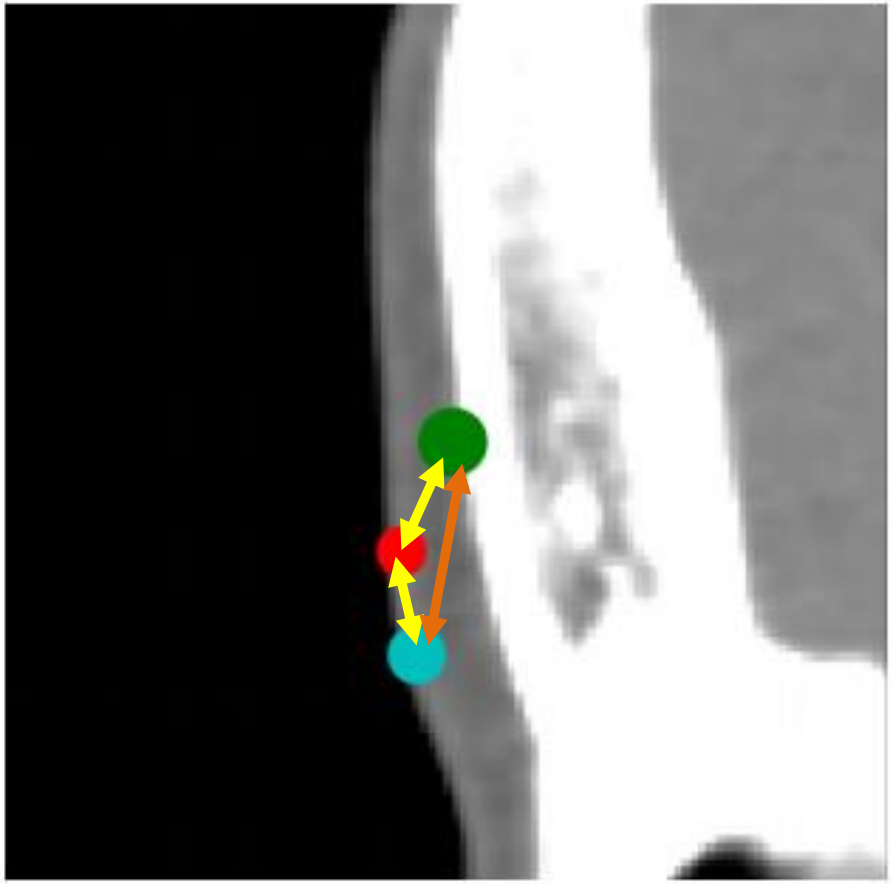


[O'Neil et al. 2018]

Metric: How far were we from the landmark in mm

When can we say we're doing as well as a radiologist?
1mm away?
0.5mm away?

How good is good enough?



[O'Neil et al. 2018]

Metric: How far were we from the landmark in mm

Green and blue are 2 different specialist annotators, red is the model.

How good is good enough?

- To be as good as a difference between 2 humans completing the same task

Method	Reference							
	Observer A				Observer B			
	Mean	Median	Max	%	Mean	Median	Max	%
Observer A	-	-	-	-	2.20	1.49	9.27	11.0
Observer B	2.20	1.48	9.27	11.0	-	-	-	-
Pass 0 (4mm)								
DF	4.47	4.03	11.54	50.4	4.58	4.14	11.28	49.2
DF (+HOG)	4.25	3.91	10.07	47.7	4.36	3.92	9.86	46.5
FCN	3.38	2.65	12.20	21.6	3.52	2.71	12.15	23.7
FCN + Atlas Correction	3.03	2.53	10.45	16.1	3.31	2.62	10.89	21.6
Pass 1 (2mm)								
DF	3.59	2.85	13.73	26.6	3.83	3.02	13.59	27.6
DF (+HOG)	3.30	2.88	**9.77**	24.9	3.47	2.84	**9.69**	26.9
FCN	2.93	1.50	19.08	12.2	3.42	1.84	20.93	17.5
FCN + Atlas Correction	**2.29**	**1.49**	11.41	10.8	**2.77**	**1.78**	12.26	**16.1**
Alternative: Pass 0 (2mm) + Atlas Correction								
FCN	2.55	1.55	15.08	**10.3**	3.10	1.92	15.38	17.5

- Read further details in the paper:

Alison Q. O'Neil, Antanas Kascenas, Joseph Henry, Daniel Wyeth, Matthew Shepherd, Erin Beveridge, Lauren Clunie, Carrie Sansom, Evelina Seduikyte Keith Muir, Ian Poole (2018).

Attaining human-level performance with atlas location autocontext for anatomical landmark detection in 3D CT data

The European Conference on Computer Vision (ECCV) Workshops, 2018

Which metric?

- Multi-label classification task (non-exclusive labels)

Lorem ipsum dolor sit amet, consectetur adipiscing elit. Aenean malesuada velit finibus nisi vulputate posuere. Proin finibus eros condimentum commodo condimentum. Integer nec efficitur augue, a gravida diam. Fusce convallis quam eu ex vehicula, nec feugiat enim dignissim. Etiam vitae felis quis enim varius scelerisque id a augue. Donec ultrices maximus ante. Sed porta, neque nec auctor facilisis, libero ligula hendrerit purus, sit amet volutpat arcu orci et arcu. Nunc maximus commodo sapien, non eleifend dolor egestas ac. Suspendisse sed arcu eget nisi pharetra tempus vel sed erat. Mauris eleifend interdum augue eget viverra.

Praesent malesuada leo at sapien tincidunt finibus. Duis tincidunt et justo at pellentesque. Sed porta egestas pretium. Fusce vestibulum fermentum quam, ornare sollicitudin sapien volutpat in. Vestibulum ultrices accumsan lorem, ut euismod leo vulputate at. Vivamus ex mauris, scelerisque eget dui eu, pharetra fermentum nulla. Class aptent taciti sociosqu ad litora torquent per conubia nostra, per inceptos himenaeos. Nam vitae ultrices augue. [illegible] erat finibus pellentesque dui. Nunc eget erat non ex volutpat hendrerit. Curabitur scelerisque [illegible] nec ligula blandit pulvinar. Integer maximus in diam sed volutpat. Sed maximus eget dui id aliquet. Suspendisse elit eros.

Hospital discharge summary

Integer mollis turpis id molestie tincidunt. Morbi et dictum turpis. Mauris eleifend dui a odio accumsan sollicitudin. Suspendisse semper dui sit amet pellentesque posuere. [illegible] felis, consequat faucibus odio ullamcorper et. Nulla scelerisque sagittis sapien.

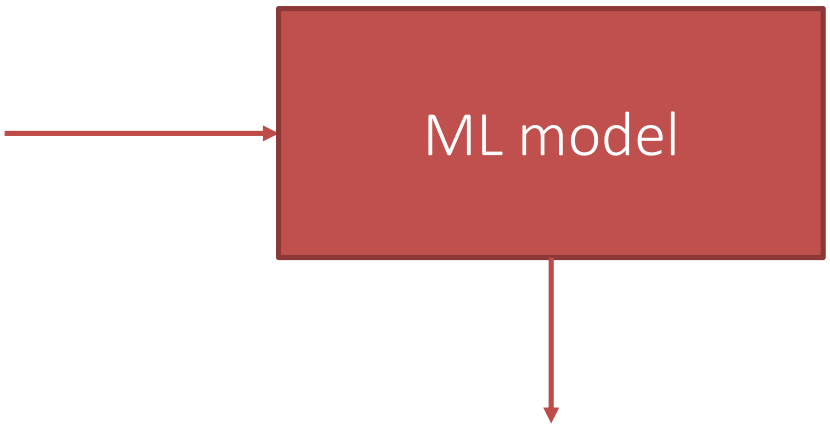
Phasellus at elit sed libero lobortis venenatis quis lacinia dolor. Cras quis massa ex. Nam ac orci non elit mattis porttitor. Vivamus venenatis elit purus, ullamcorper sodales turpis malesuada quis. Nullam sapien nulla, ullamcorper sit amet euismod eget, aliquam sit amet nulla. Vivamus a consequat felis. [illegible] massa eleifend, quis euismod erat laoreet. In hac habitasse platea dictumst. Sed et pu[illegible] vehicula vitae vehicula quis, sodales vel ante. Quisque vestibulum, magna sed euismod aliquam, nisi mauris finibus lacus, a posuere justo arcu id est. Aliquam ultricies pretium nibh et porta. Vivamus dictum, libero sit amet egestas ullamcorper, eros metus posuere dui, a lobortis nisi erat quis sapien. Nullam sit amet est eu est tristique pretium. Quisque pellentesque consequat eros eu imperdiet.

Maecenas consectetur ex sit amet leo dictum, id blandit orci dapibus. Pellentesque ut risus condimentum, porta augue quis, placerat nunc. Aenean vitae ipsum ultrices, elementum ex id, facilisis arcu. Praesent eu eros eleifend, sollicitudin urna nec, aliquam massa. Vestibulum eros urna, tincidunt eget est vitae, efficitur finibus libero. Curabitur varius, dui id porta porttitor, mauris mauris convallis ipsum, commodo rutrum nulla diam sit amet dolor. Quisque congue faucibus sem, nec facilisis orci maximus et.

Donec ut dapibus tortor, a varius nisi. Ut interdum, enim id maximus vehicula, enim mauris pharetra dui, in mollis urna urna vel erat. Suspendisse vehicula nunc non placerat iaculis. Quisque interdum metus risus, id venenatis odio mollis ac. Praesent at erat dui. Donec nulla leo, pellentesque vitae sapien tempor, pharetra dictum dui. Donec vitae luctus lorem. Praesent porta, lorem et luctus sagittis, est nulla pretium eros, id tincidunt lectus orci sit amet nisl. Sed vel imperdiet turpis. Vestibulum ante ipsum primis in faucibus orci luctus et ultrices posuere cubilia Curae; Nunc arcu dui, tincidunt eget porta in, bibendum iaculis metus. Cras ultricies lectus velit, non cursus urna venenatis vel. Donec ultricies at eros a vestibulum. Pellentesque nisl lectus, viverra placerat lorem quis, ornare vehicula eros. Integer et quam id velit efficitur bibendum. Aenean molestie elit nibh, in accumsan ante vestibulum in.

Duis mollis elementum justo, in ornare neque porttitor ultrices. Sed tincidunt justo augue. Integer venenatis iaculis nisi, vel porttitor quam feugiat id. Sed velit quam, tristique ut turpis in, vehicula blandit est. Morbi in fringilla mi, ac iaculis sem. Duis consequat accumsan tortor vitae malesuada. Orci varius natoque penatibus et magnis dis parturient montes, nascetur ridiculus mus. Vivamus rhoncus eros purus, quis rutrum sem tempus eget. Duis interdum pretium nisl, in pharetra quam porttitor ut. Suspendisse a sem a ante faucibus bibendum ut eu ex. Class aptent taciti sociosqu ad litora torquent per conubia nostra, per inceptos himenaeos. Nunc pellentesque enim non metus hendrerit pellentesque. Quisque pellentesque enim lacus, et congue neque tempus id. Curabitur eu neque eget sem faucibus accumsan. Nam ac vulputate dolor.

Sed neque neque, iaculis quis porttitor sit amet, ultrices nec purus. Praesent risus nulla, egestas id scelerisque sed, aliquam vel diam. Aenean hendrerit vel tortor vel laoreet. In aliquam leo eget dignissim condimentum. Vestibulum facilisis pulvinar feugiat. Nullam sit amet lacinia orci. Maecenas a lacus et orci vestibulum dignissim et quis ipsum. Phasellus vel maximus nisl. Vestibulum tincidunt tempus nisi, quis mollis arcu mollis vitae. Duis molestie pellentesque dui porttitor aliquet. Etiam ut euismod dui, quis iaculis quam.



	hypertension	lung cancer	diabetes	coronary artery disease
John	1	0	1	1
Ewan	0	1	1	0

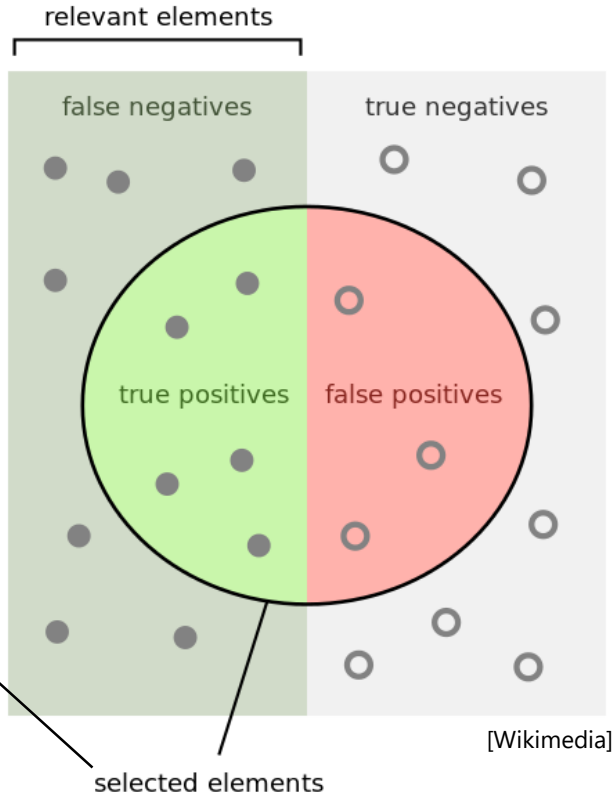
~9,000 labels, 10-50 ‘1’s per patient

Canon

Which metric?

- Predictions for one disease:

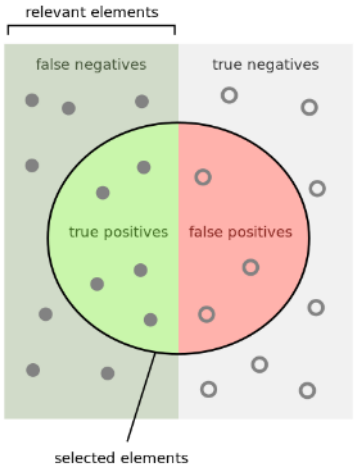
	True	Predicted
Thomas	1	0.95
Caroline	1	0.88
Sarah	0	0.65
Mike	1	0.62
Dave	1	0.48
Michelle	0	0.43
Anna	0	0.26
Alex	0	0.25
George	1	0.22
Emily	0	0.03
Kate	0	0.01



[Wikimedia]

Which metric?

- Accuracy + AUC (Area under the ROC curve)

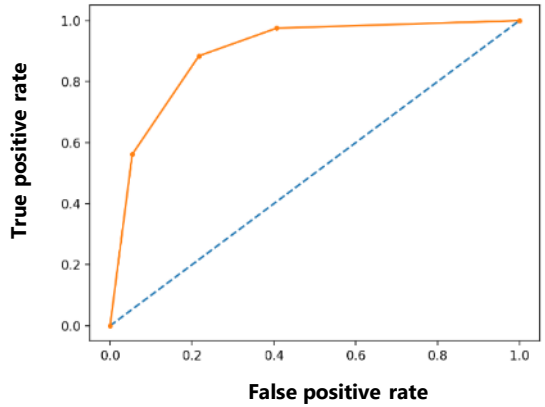


Accuracy

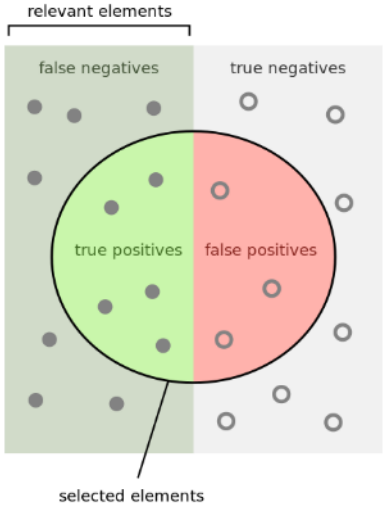
Recall
Sensitivity
True positive rate
Hit rate

Specificity
True negative rate

Fall-out
False positive rate



- Precision, Recall, F1 + AUPRC (Area under the precision-recall curve)



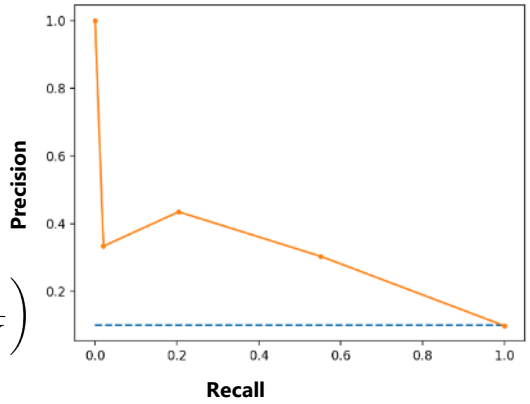
How many selected items are relevant?

Precision
Positive predictive value

How many relevant items are selected?

Recall
Sensitivity
True positive rate
Hit rate

$$F_1 = \left(\frac{\mathrm{recall}^{-1} + \mathrm{precision}^{-1}}{2}\right)^{-1} = \left(\frac{2TP}{2TP + FP + FN}\right)$$



Canon

How to average

	hypertension	coronary artery disease	…	Jaw dislocation	Struck by a lightning
John	1 0.89	0	…	1	0
Ewan	1 0.78	1	…	0	0
Alice	0 0.12	0		0	0
…					
Total cases	10,000	3,000	…	10	2

How to average

	hypertension	coronary artery disease	…	Jaw dislocation	Struck by a lightning
John	1 0.89 TP	0	…	1	0
Ewan	1 0.78 TP	1	…	0	0
Alice	0 0.12 TN	0		0	0
…					
Total cases	10,000	3,000	…	10	2
F1 score	0.85	0.7		0.2	0.0

true prediction

$$F_1(hypertension) = \left(\frac{\text{recall}^{-1}+\text{precision}^{-1}}{2}\right)^{-1} = \left(\frac{2TP}{2TP+FP+FN}\right) = 0.85$$

How to average

	hypertension	coronary artery disease	…	Jaw dislocation	Struck by a lightning
Total cases	10,000	3,000	…	10	2
F1 score	0.85	0.7		0.2	0.0

$$F_1(all) = \frac{0.85+0.7+\ldots+0.2+0}{N_{labels}} = 0.088$$

← [Mullenbach et al. 2018]

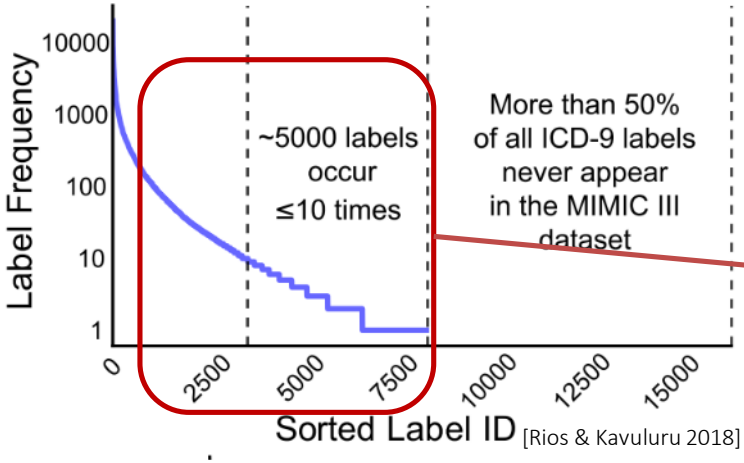
How to average

	hypertension	coronary artery disease	...	Jaw dislocation	Struck by a lightning
Total cases	10,000	3,000	...	10	2
F1 score	0.85	0.7		0.2	0.0

$$F_1(all) = \frac{0.85+0.7+\ldots+0.2+0}{N_{labels}} = 0.088$$

[Mullenbach et al. 2018]

Hard to perform well on rare labels



[Rios & Kavuluru 2018]

Majority of labels are rare (<100 positive examples)

How to average

	hypertension	coronary artery disease	...	Jaw dislocation	Struck by a lightning
John	1 0.89 TP	0 0.55 FP	...	1 0.75 TP	0 0.01 TN
Ewan	1 0.78 TP	1 0.48 FN	...	0 0.34 TN	0 0.02 TN
Alice	0 0.12 TN	0 0.05 TN		0 0.62 FP	0 0.01 TN
...					
Total cases	10,000	3,000	...	10	2

true prediction

How to average

	hypertension	coronary artery disease	...	Jaw dislocation	Struck by a lightning
John	1 0.89 TP	0 0.55 FP	...	1 0.75 TP	0 0.01 TN
Ewan	1 0.78 TP	1 0.48 FN	...	0 0.34 TN	0 0.02 TN
Alice	0 0.12 TN	0 0.05 TN		0 0.62 FP	0 0.01 TN
...					
Total cases	10,000	3,000	...	10	2

true prediction

$$F_1 = \left(\frac{\text{recall}^{-1} + \text{precision}^{-1}}{2} \right)^{-1} = \left(\frac{2total_{TP}}{2total_{TP} + total_{FP} + total_{FN}} \right) = 0.54$$

[Mullenbach et al. 2018]

How to average

	hypertension	coronary artery disease	...	Jaw dislocation	Struck by a lightning
John	1 0.89 TP	0 0.55 FP	...	1 0.75 TP	0 0.01 TN
Ewan	1 0.78 TP	1 0.48 FN	...	0 0.34 TN	0 0.02 TN
Alice	0 0.12 TN	0 0.05 TN		0 0.62 FP	0 0.01 TN
...					
Total cases	10,000	3,000	...	10	2

true prediction

$$F_1 = \left(\frac{\text{recall}^{-1} + \text{precision}^{-1}}{2} \right)^{-1} = \left(\frac{2total_{TP}}{2total_{TP} + total_{FP} + total_{FN}} \right) = 0.54$$

$$0.088$$

macro

[Mullenbach et al. 2018]

How to average

	hypertension	coronary artery disease	...	Jaw dislocation	Struck by a lightning
John	1 0.89 TP	0 0.55 FP	...	1 0.75 TP	0 0.01 TN
Ewan	1 0.78 TP	1 0.48 FN	...	0 0.34 TN	0 0.02 TN
Alice	0 0.12 TN	0 0.05 TN		0 0.62 FP	0 0.01 TN
...					
Total cases	10,000	3,000	...	10	2

true prediction

$$F_1 = \left(\frac{\text{recall}^{-1} + \text{precision}^{-1}}{2}\right)^{-1} = \left(\frac{2total_{TP}}{2total_{TP} + total_{FP} + total_{FN}}\right) = 0.54$$

0.088 macro

0.54 micro

[Mullenbach et al. 2018]

Which one to choose

- Score for each label then average over all labels – **macro average**
- Treat all predictions equally, calculate a single score – **micro average**

Which one to choose

- Score for each label then average over all labels – **macro average**
- Treat all predictions equally, calculate a single score – **micro average**

- Arguments:
  - Is per-label prediction the definition of the task? Or per patient?
  - Distribution of diagnosis prevalence similar between training and deployment
  - Cost of misclassification variable between diseases
  - Impact on perception
  - Direction of work on the model

Which one to choose

- Score for each label then average over all labels – **macro average**
- Treat all predictions equally, calculate a single score – **micro average**

- Arguments:
  - Is per-label prediction the definition of the task? Or per patient?
  - Distribution of diagnosis prevalence similar between training and deployment
  - Cost of misclassification variable between diseases
  - Impact on perception
  - Direction of work on the model

- Maybe we should've picked accuracy instead…

Conclusion

- Think about the use case!
- (here we don't know how good/consistent human annotators are)
- Want to learn more about state-of-the-art deep learning solutions for ICD coding of hospital discharge summaries:
  - Mullenbach, J., Wiegreffe, S., Duke, J., Sun, J., & Eisenstein, J. (2018). **Explainable Prediction of Medical Codes from Clinical Text.** *Proceedings Of NAACL-HLT 2018*, 1101–1111.

Made For life

For over 100 years, the Canon Medical Systems `Made for Life' philosophy prevails as our ongoing commitment to humanity - generations of inherited passion creates a legacy of medical innovation and service that continues to evolve as we do. By engaging the brilliant minds of many, we continue to set the benchmark, because we believe quality of life should be a given, not the exception.

Canon

CANON MEDICAL SYSTEMS