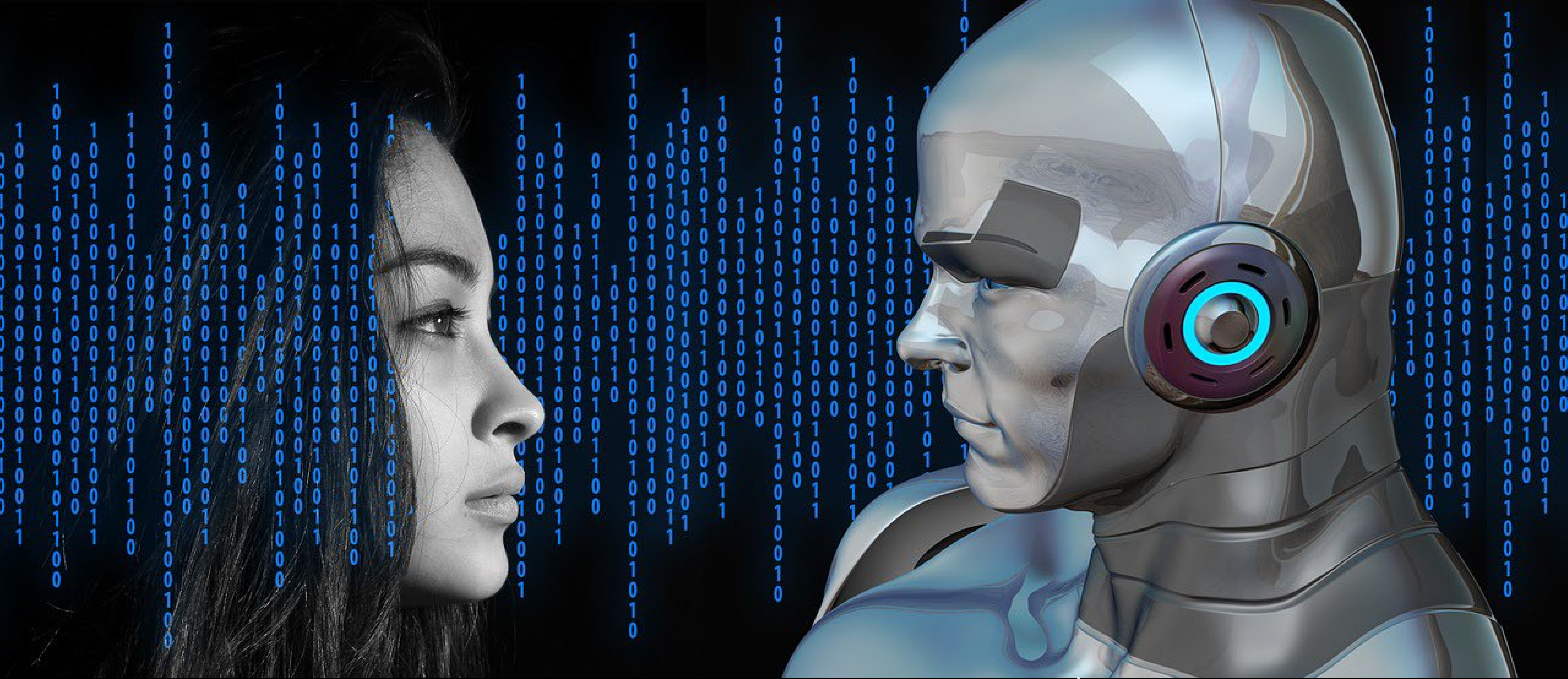
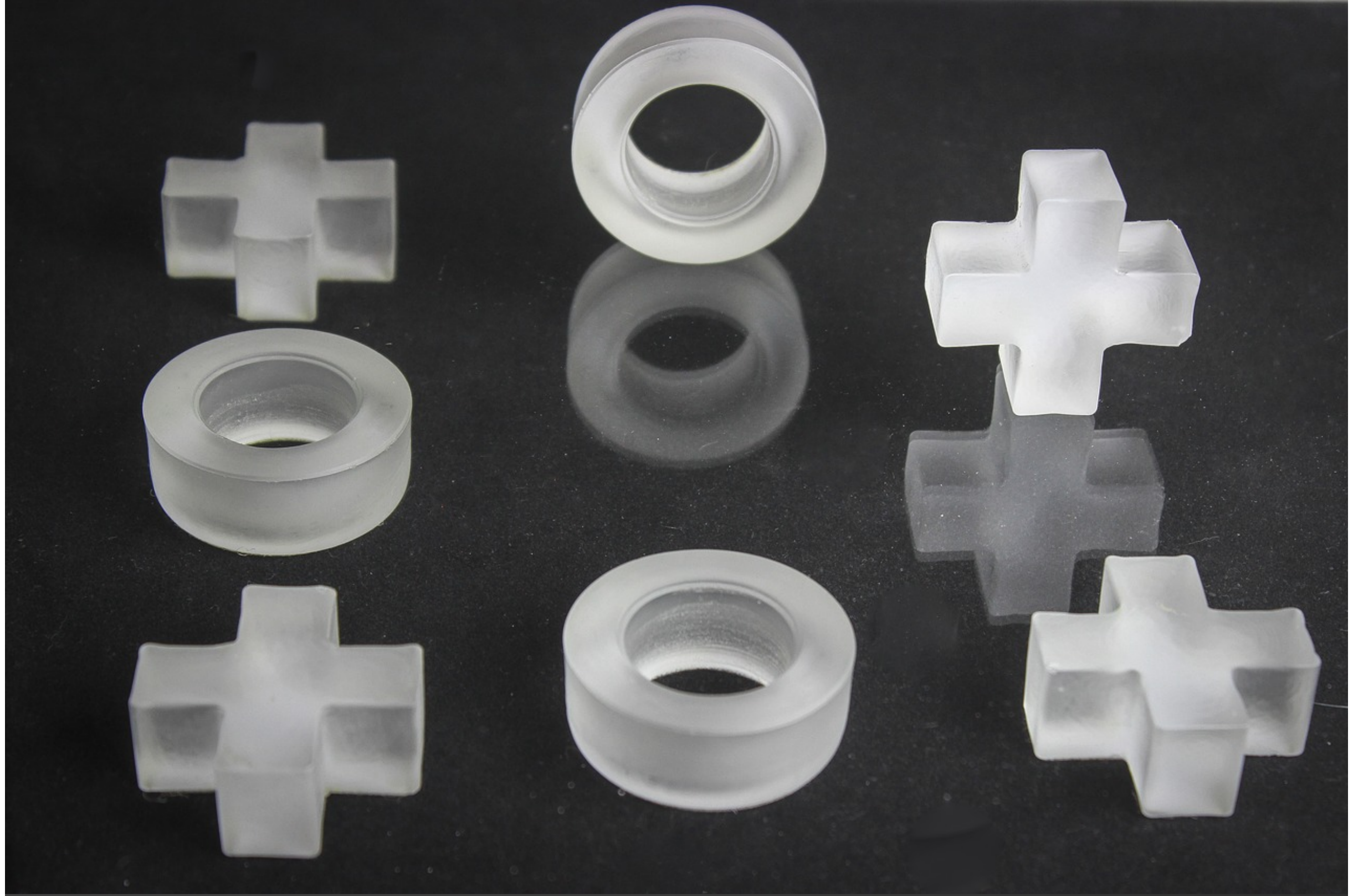


The Biomedical AI Revolution

David C. Anastasiu
Assistant Professor
San José State University

According to Miriam-Webster:

1. A branch of computer science dealing with the simulation of intelligent behavior in computers.
2. The capability of a machine to imitate intelligent human behavior.



What is AI?

Types of AI

Narrow	General	Augmented
Pattern Recognition Prediction	Can think like a human	Can help a human think

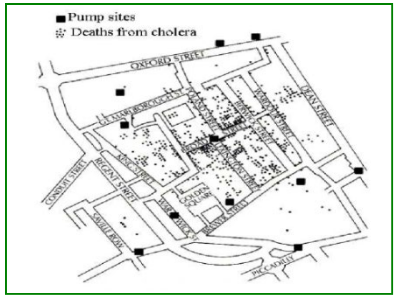
Why Care About AI?

- There has been enormous data growth in both commercial and scientific databases due to advances in data generation and collection technologies

- New mantra
  - Gather whatever data you can whenever and wherever possible.

- Machine learning is becoming ubiquitous in society
  - Here's a movie you might like (*Recommender Systems*)
  - You should try this product (*Advertising*)
  - Found a shortcut that will save you 10 minutes (*Maps*)
  - Jessica is a friend suggestion for you (*Social Media*)
  - Caution! Vehicle approaching in your right lane (*Drive Assist / Self-Driving Cars*)

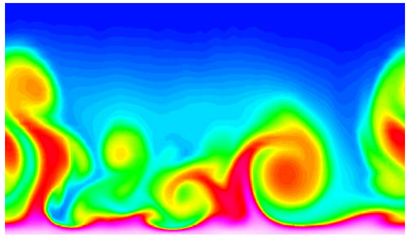
Homeland Security



Geo-spatial data

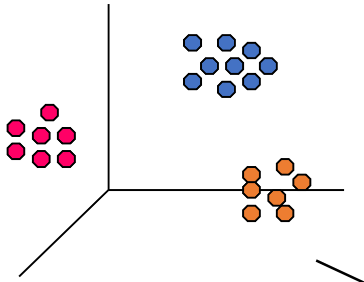
Business Data

Sensor Networks



Computational Simulations

Machine Learning Tasks

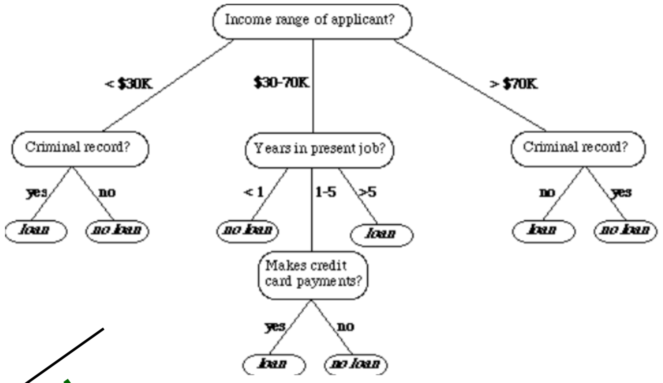


Clustering

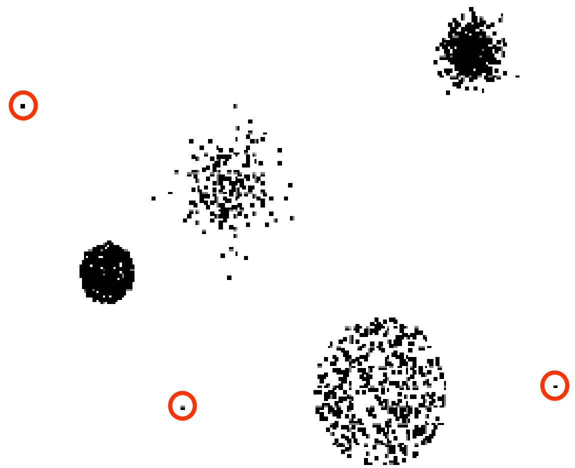
Data

Tid	Refund	Marital Status	Taxable Income	Cheat
1	Yes	Single	125K	No
2	No	Married	100K	No
3	No	Single	70K	No
4	Yes	Married	120K	No
5	No	Divorced	95K	Yes
6	No	Married	60K	No
7	Yes	Divorced	220K	No
8	No	Single	85K	Yes
9	No	Married	75K	No
10	No	Single	90K	Yes
11	No	Married	60K	No
12	Yes	Divorced	220K	No
13	No	Single	85K	Yes
14	No	Married	75K	No
15	No	Single	90K	Yes

Predictive Modeling



Anomaly Detection



Association Rules

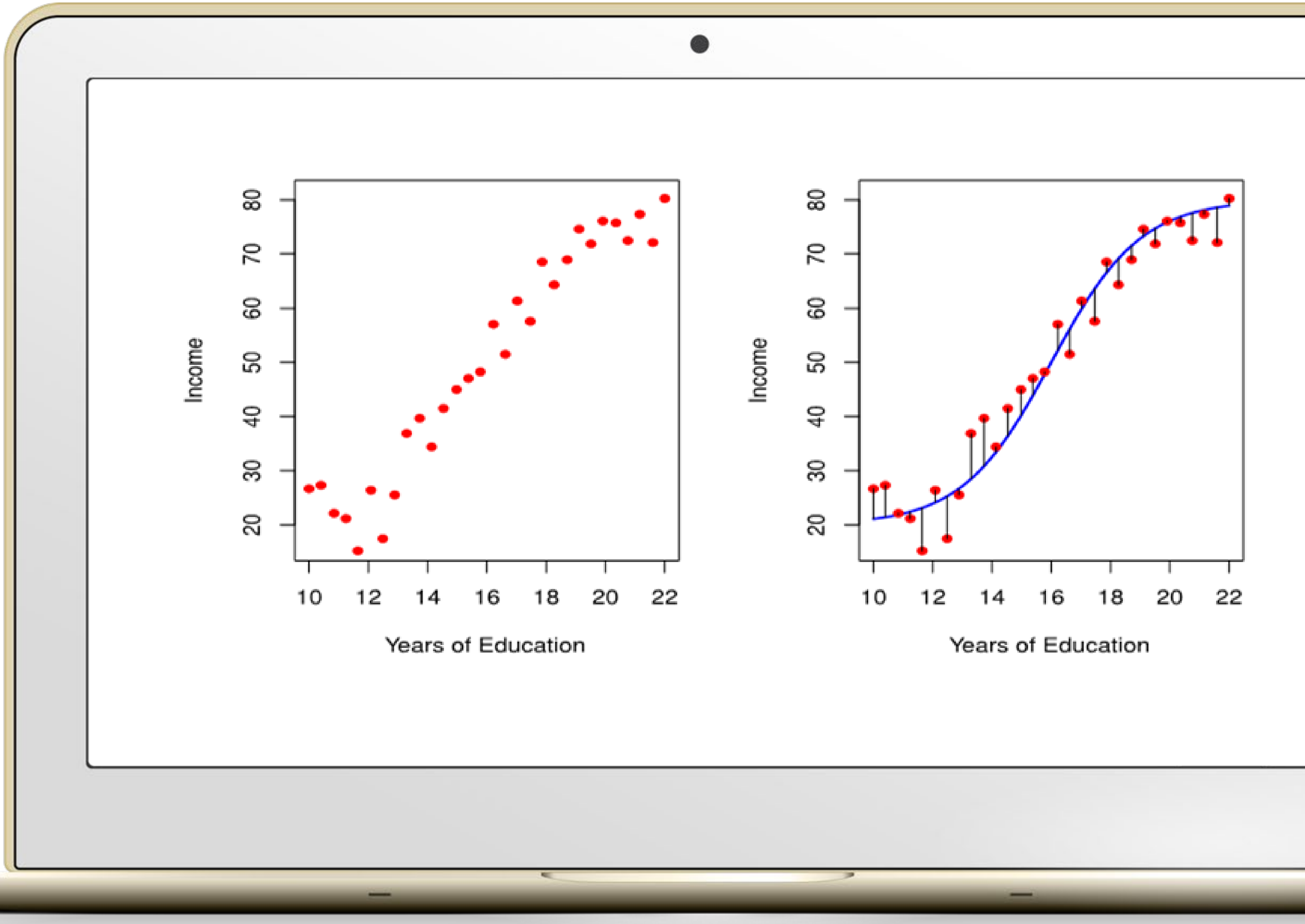
Supervised Learning

Learn a function

$$Y = f(X) + \epsilon$$

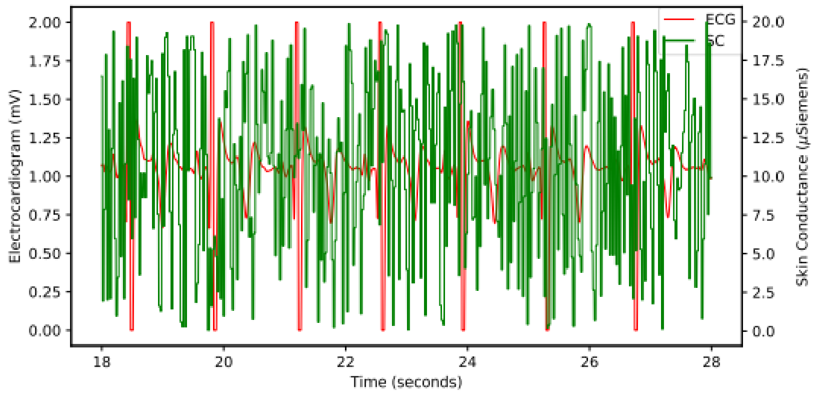
Fundamental Assumption of Learning:

- The distribution of training examples is identical to the distribution of test examples (including future unseen examples).
- Training examples must be sufficiently representative of (future) test data.



Time Series Analysis

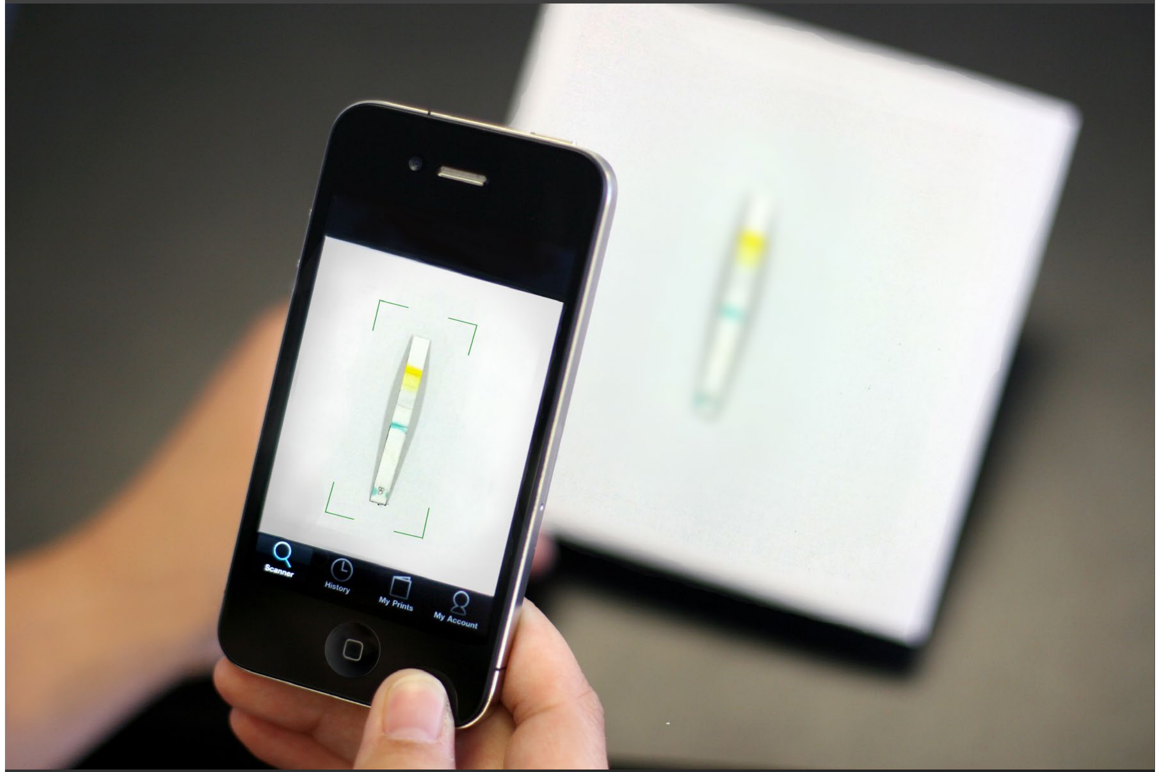
- Data collected during a sensory challenge protocol (SCP) in which the reactions to eight stimuli were observed.
- Based on electrocardiogram (ECG) and skin conductance.
- Multivariate time series /w 2M+ samples for each subject.



Example:

Kidney Health Monitoring

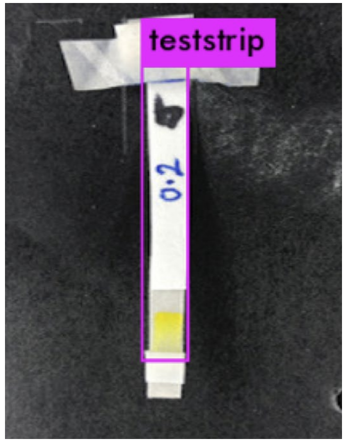
- /w Dr. Megan C. Chang
- Student: Manika Kapoor

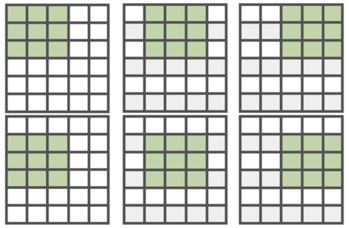


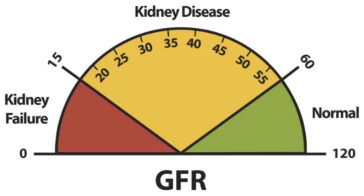
Example:
Kidney Health Monitoring

/w Dr. Alessandro Bellofiore

Students: Rathna Ramesh, Ragwa Elsayed



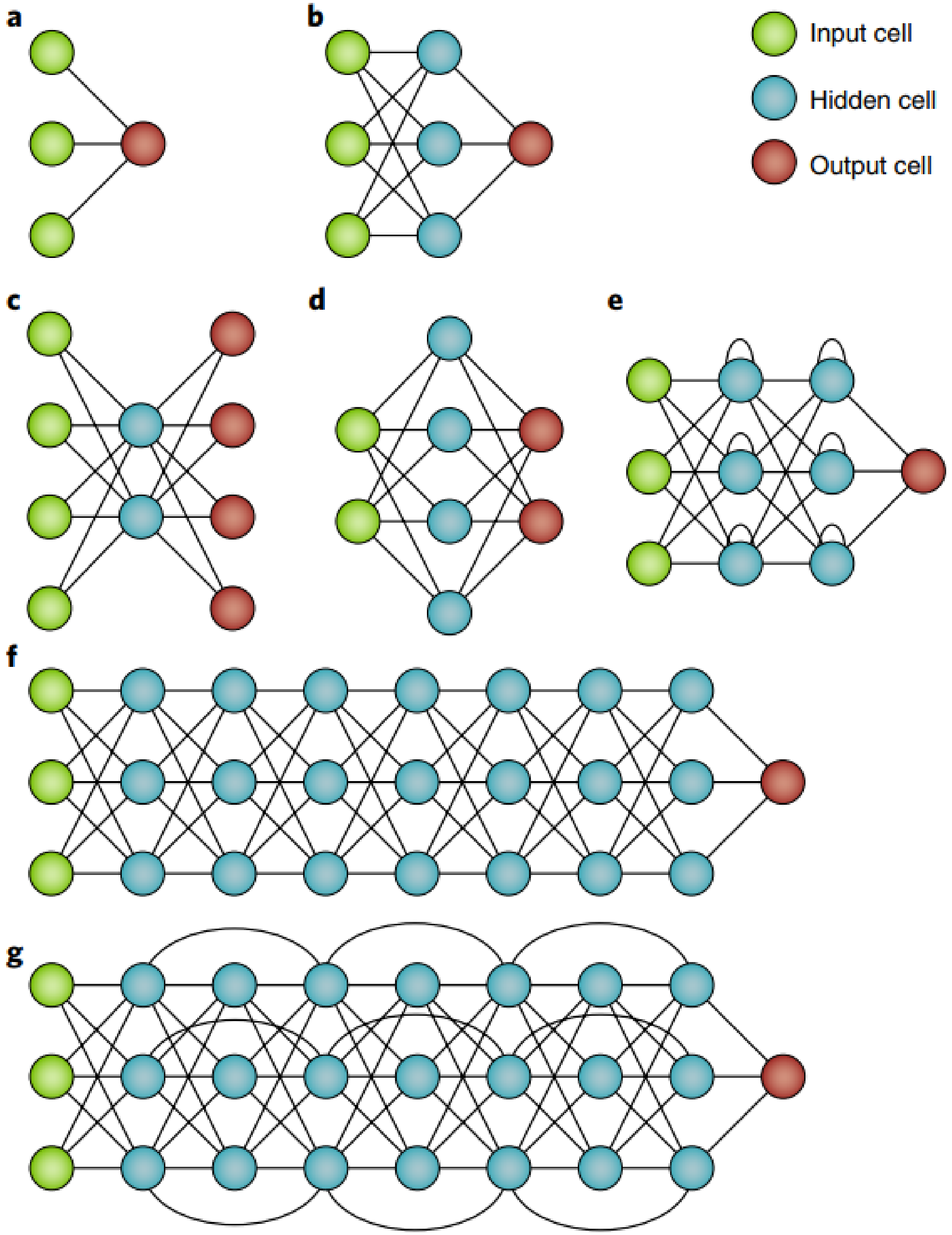




Localization	Feature Extraction	Prediction
Deep Learning – YOLO Alternatives	Color-based features: RGB, Color Histogram, Gradient Histogram	Creatinine level (regression) Kidney health level (classification)

Deep Learning

- A subfield of machine learning.
- Uses artificial neural networks (ANNs) with many layers for pattern discovery.
- Building block: Perceptron
  - $\boldsymbol{w} \cdot \boldsymbol{x} + b > 0$
- DNNs can approximate infinite functions
- Needs sufficient labeled input
  - Avoid overfitting



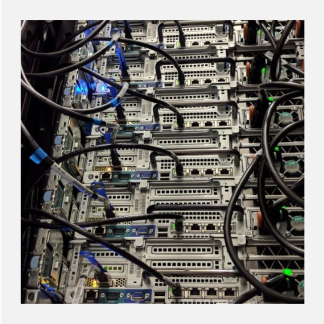
Credit: Figure adapted from Yu et al., Artificial Intelligence in healthcare, Nature Biomedical Engineering, VOL 2, Oct. 2018, 719–731, https://www.nature.com/articles/s41551-018-0305-z.pdf

What Made Deep Learning Possible

Many Core Hardware

GPUs, TPUs

Thousands of cores, really good at dense matrix operations.



Distributed Computing

Supercomputing
Shared-nothing Computing

Split the work among many systems.

Lots of Labelled Data

Mechanical Turk
Label Generation

Split the work among many humans, or be clever about creating labels

Example: Labeled Data Generation

2017 AI City Challenge

- Collaborative annotation
- Over 150,000 annotations from 80 h video
- Localization and classification

https://youtu.be/Jlvh_KxHl40 Thomas Tang & team - UW

https://youtu.be/mtu9_w8B984 Charles MacKay & team - SJSU

2018 AI City Challenge

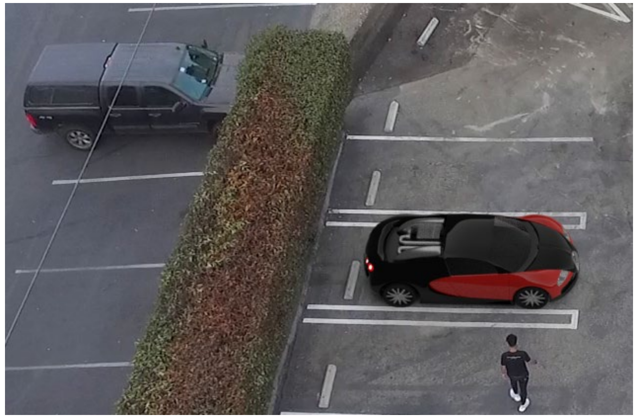
- GPS-based annotation
- 27 videos, 3 locations
- Speed estimation, anomaly detection, re-identification and tracking

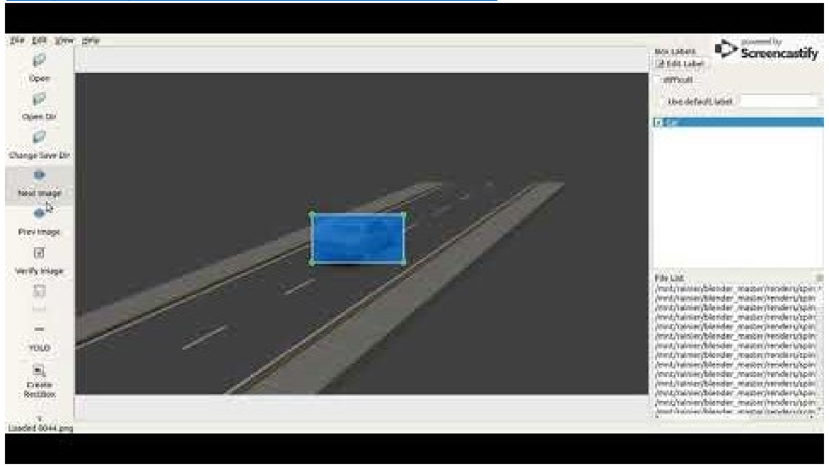
Example: Labeled Data Generation

https://youtu.be/3tnCaMBVT8Q Charles MacKay

CGI-based Labelling

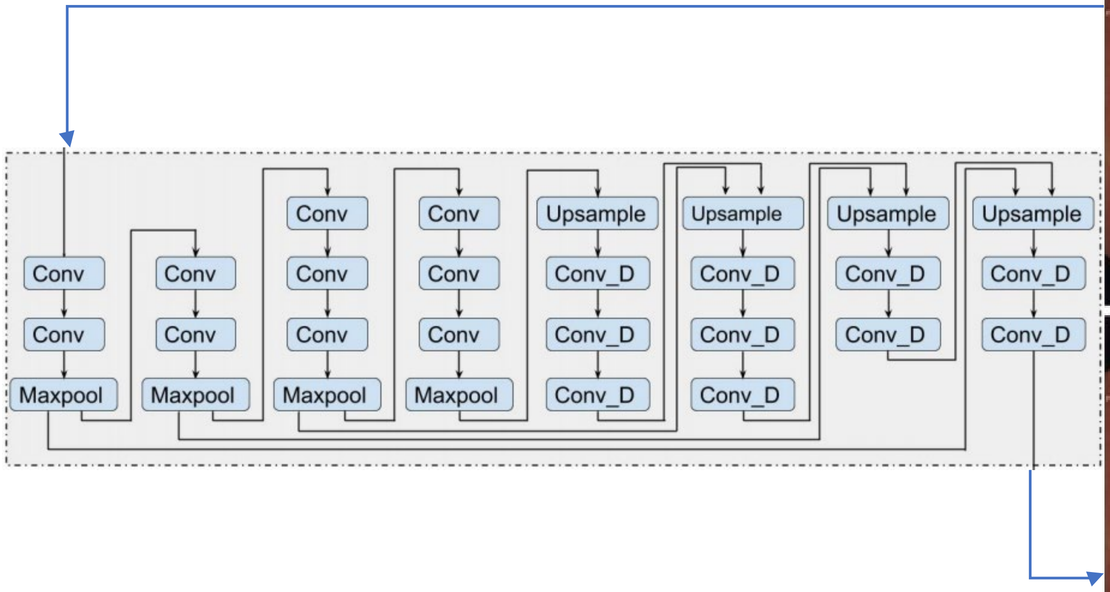
- Blender-generated objects w/ motion
- Bounding box automatically generated
- GAN-based smoothing to improve blend

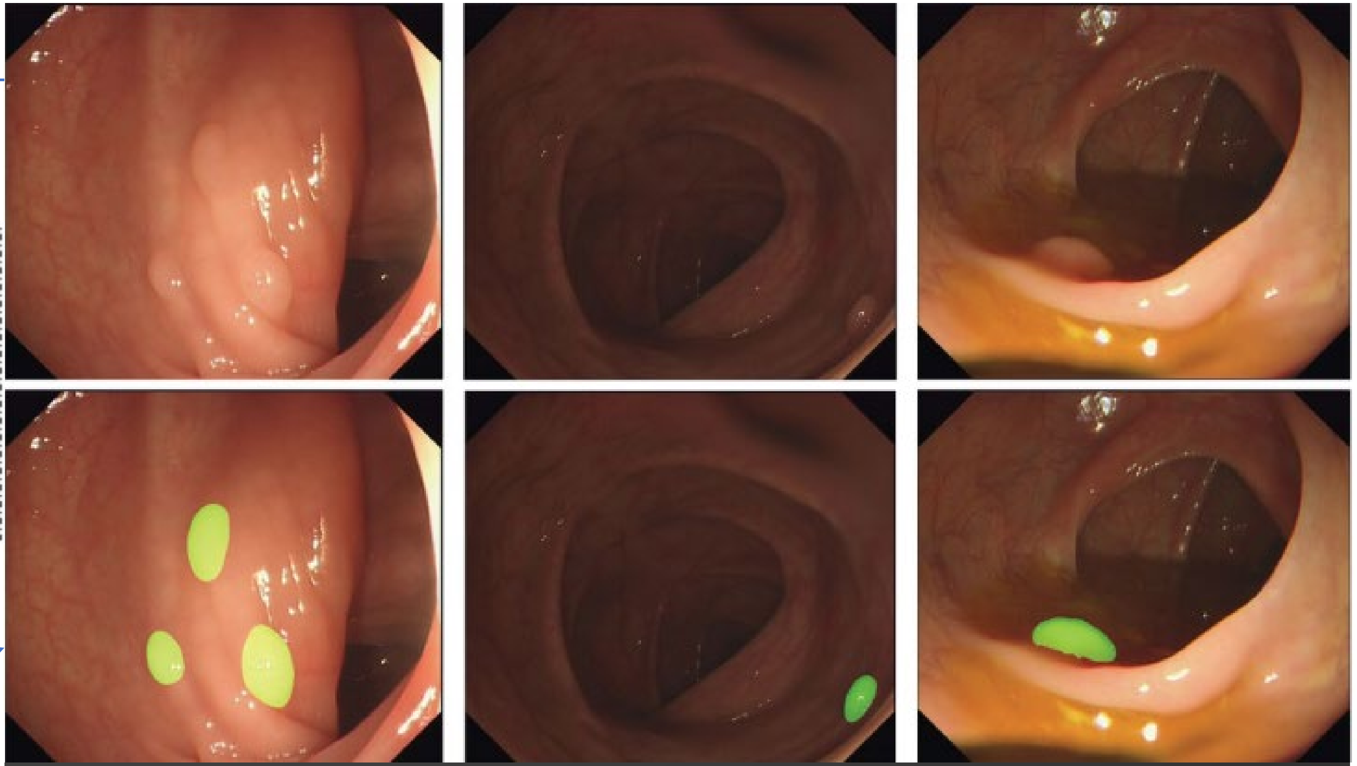




Anomaly Detection in Expense Reports

- Receipt localization dataset generation
- Small representative set of receipts
- GAN-based model for background
- Generate millions of receipts & combinations





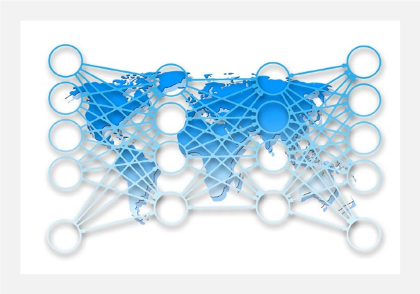
Deep Learning

- Convolutional Neural Network (CNN)
- SegNet Architecture
- Computer probability of each pixel belonging to a polyp

Example: **Real-time detection of polyps during colonoscopy using machine learning.**

Credit: Figures adapted from Wang et al., Development and validation of a deep-learning algorithm for the detection of polyps during colonoscopy. Nature Biomedical Engineering, volume 2, 741–748 (2018) https://www.nature.com/articles/s41551-018-0315-x.pdf

The Challenges of Debugging Deep-Learning Software



Data Driven Behaviour

GPUs, TPUs

What is the path that will be taken at inference time?

Input Bias

(Un)intended consequences

Who’s being left out?

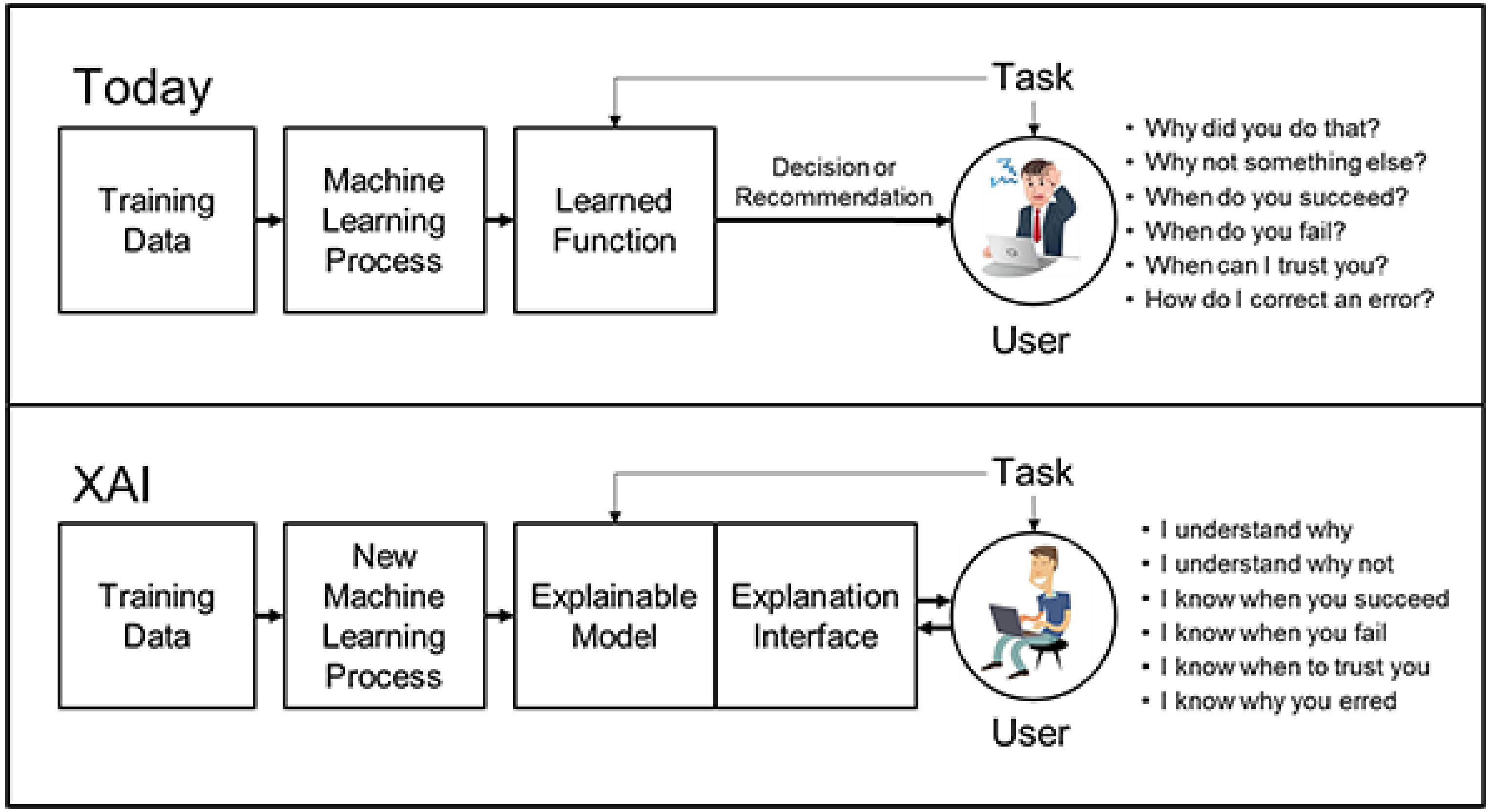
XAI

Explainable AI

Understand why the model made the decision

Explainable AI

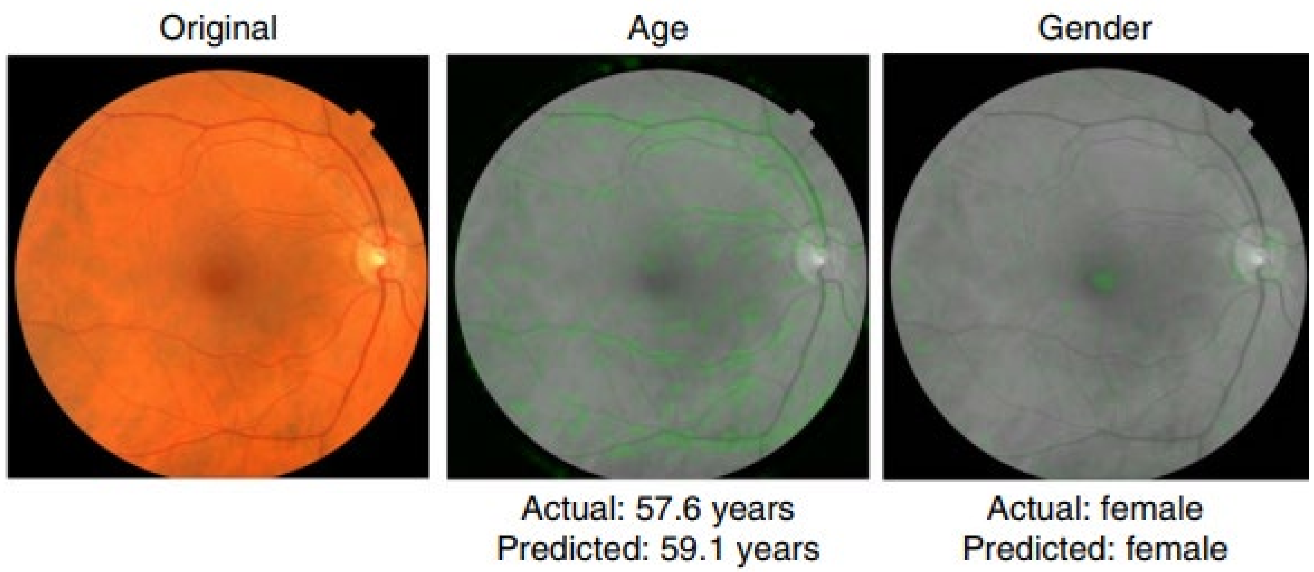
- Critical requirement for AI in many contexts beyond healthcare
  - DARPA/DoD priority
  - Already required in the General Data Protection Right (GDPR)
- RISE: randomized input sampling for explanation of black-box models



Credit: Figure adapted from David Gunning, Explainable Artificial Intelligence (XAI), https://www.darpa.mil/program/explainable-artificial-intelligence

Deep Learning

- Convolutional Neural Network (CNN)
- Improved prediction of cardiovascular risk factors
- Predicted side-factors:
  - Age, gender, smoking, diabetic, BMI
- “Soft attention” used to point out the salient pixels

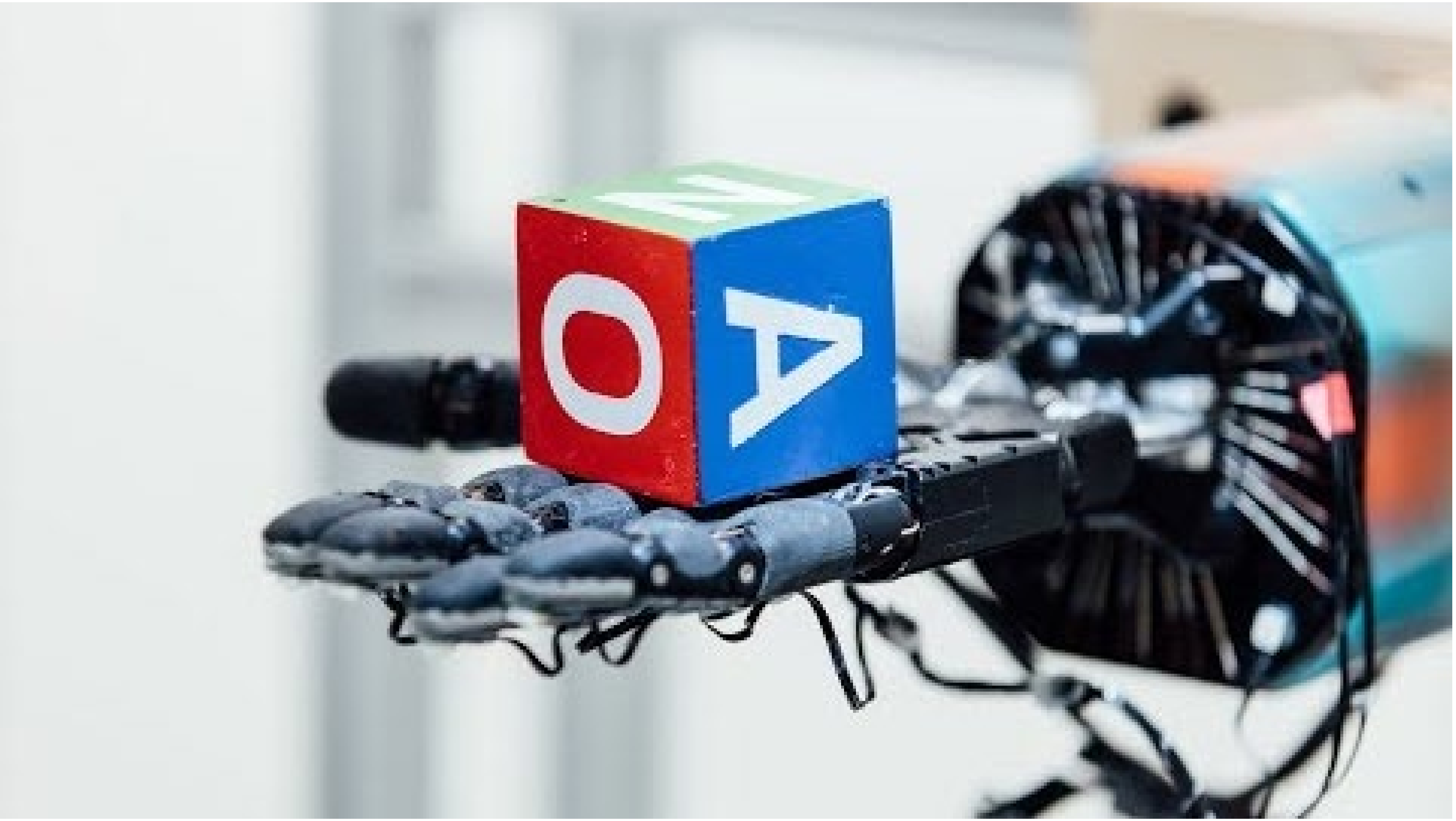


Example: **Prediction of cardiovascular risk factors from retinal fundus photographs via deep learning**

Credit: Figures adapted from Poplin et al., Nature Biomedical Engineering, volume 2, pages 158–164 (2018) https://www.nature.com/articles/s41551-018-0195-0.pdf

Reinforcement Learning

https://youtu.be/jwSbzNHGflM?mute=1



Conclusions

- AI is here to stay (this time)
- Will permeate Biomedical Engineering (and many other fields)
- Learn machine learning

  or partner with a Data Scientist

References

[1] Manika Kapoor and David C. Anastasiu. A data-driven approach for detecting autism spectrum disorders. In Peter Haber, Thomas Lampoltshammer, and Manfred Mayr, editors, Data Science – Analytics and Applications, iDSC 2019, Wiesbaden, 2019. Springer Fachmedien Wiesbaden.

[2] Shuai Hua and David C. Anastasiu. Effective vehicle tracking algorithm for smart traffic networks. In Thirteenth IEEE International Conference on Service-Oriented System Engineering (SOSE), SOSE 2019. IEEE, April 2019.

[3] Shuai Hua, Manika Kapoor, and David C. Anastasiu. Vehicle tracking and speed estimation from traffic videos. In 2018 IEEE Conference on Computer Vision and Pattern Recognition Workshops, CVPRW'18. IEEE, July 2018.

[4] Milind Naphade, Ming-Ching Chang, Anuj Sharma, David C. Anastasiu, Vamsi Jagarlamudi, Pranamesh Chakraborty, Tingting Huang, Shuo Wang, Ming-Yu Liu, Rama Chellappa, Jenq-Neng Hwang, and Siwei Lyu. The 2018 nvidia ai city challenge. In 2018 IEEE Conference on Computer Vision and Pattern Recognition Workshops, CVPRW'18. IEEE, July 2018.

[5] Niveditha Bhandary, Charles MacKay, Alex Richards, Ji Tong, and David C. Anastasiu. Robust classification of city roadway objects for traffic related applications. In 2017 IEEE Smart World NVIDIA AI City Challenge, SmartWorld'17, Piscataway, NJ, USA, 2017. IEEE.

[6] Milind Naphade, David C. Anastasiu, Anuj Sharma, Vamsi Jagrlamudi, Hyeron Jeon, Kaikai Liu, Ming-Ching Chang, Siwei Lyu, and Zeyu Gao. The nvidia ai city challenge. In 2017 IEEE SmartWorld Conference, SmartWorld'17, Piscataway, NJ, USA, 2017. IEEE.

[7] Konrad Paul Kording, Roozbeh Farhoodi, Ari S. Benjamin, and Joshua I. Glaser. The roles of machine learning in biomedical science. Winter Bridge on Frontiers of Engineering, 47(4), 2017.

[8] Ryan Poplin, Avinash V. Varadarajan, Katy Blumer, Yun Liu, Michael V. McConnell, Greg S. Corrado, Lily Peng, and Dale R. Webster. Prediction of cardiovascular risk factors from retinal fundus photographs via deep learning. Nature Biomedical Engi-neering, 2(3):158–164, 2018.

[9] Pu Wang, Xiao Xiao, Jeremy R. Glissen Brown, Tyler M. Berzin, Mengtian Tu, Fei Xiong, Xiao Hu, Peixi Liu, Yan Song, Di Zhang, Xue Yang, Liangping Li, Jiong He, Xin Yi, Jingjia Liu, and Xiaogang Liu. Development and validation of a deep-learning algorithm for the detection of polyps during colonoscopy. Nature Biomedical Engineering, 2(10):741–748, 2018.

[10] Kun-Hsing Yu, Andrew L Beam, and Isaac S Kohane. Artificial intelligence in healthcare. Nature Biomedical Engineering,2(10):719–720, 2018.

[11] Ragwa M. El Sayed, Rathna Ramesh, Alessandro Bellofiore, David C. Anastasiu, and Melinda Simon. Patient friendly kidney function screening, 2018. Poster presented by Ragwa M. El Sayed at the National Kidney Foundation 2018 Spring Clinical Meeting, Austin, TX.