



Applications and Framework

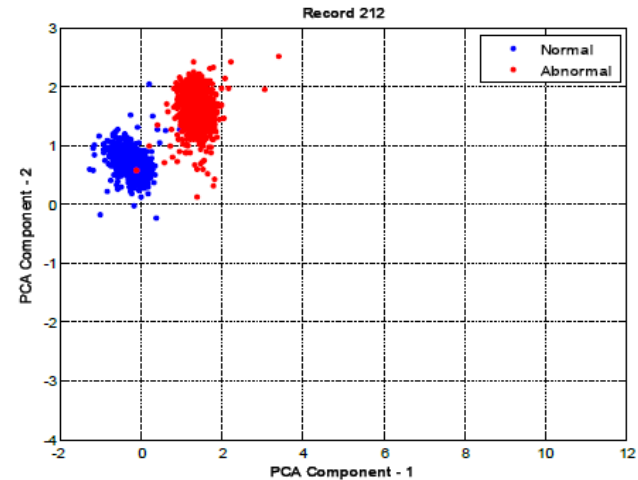
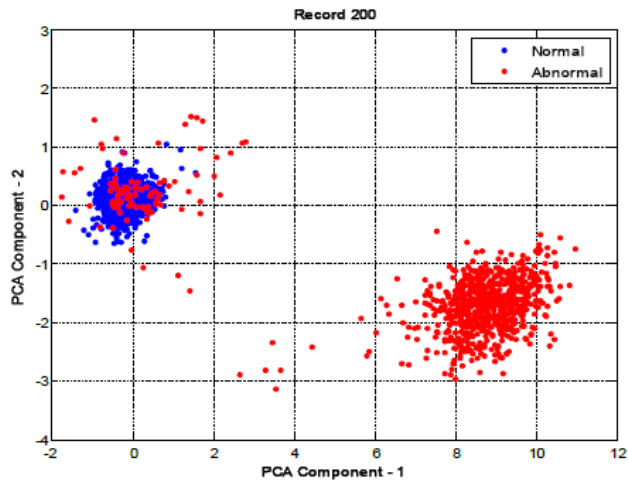
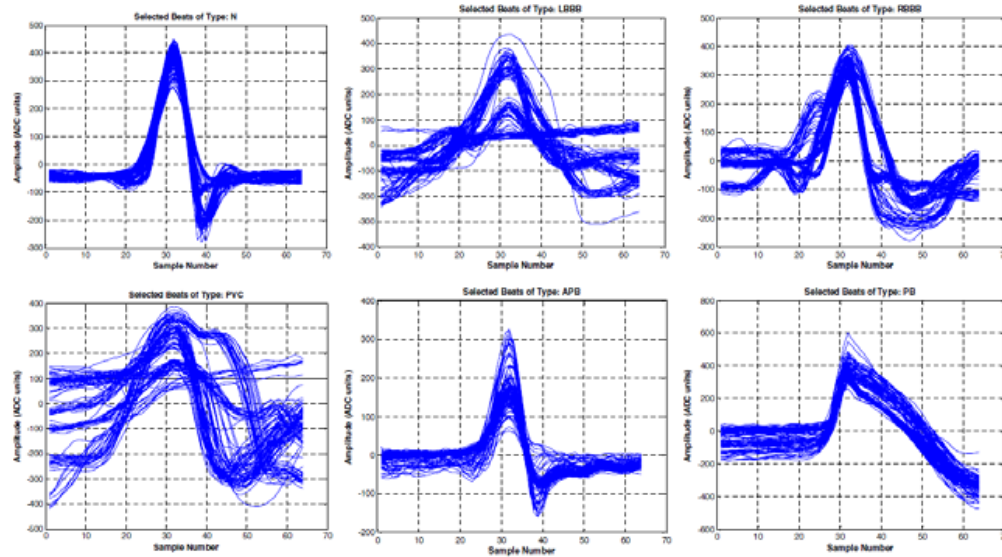
Dr. Fayyaz Minhas

Department of Computer Science
University of Warwick

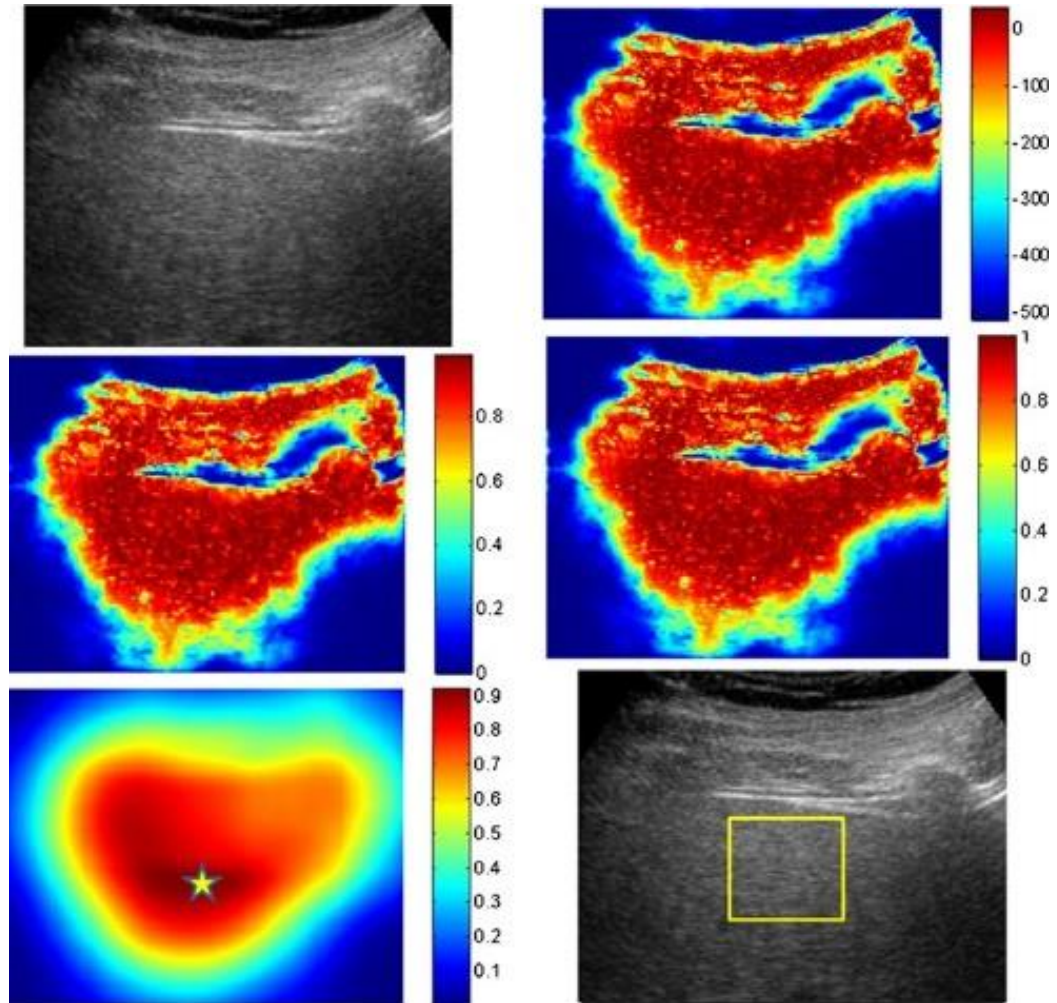
Applications

- An ability that I would like you to learn is to identify how to use data mining in different domains.

ECG Classification

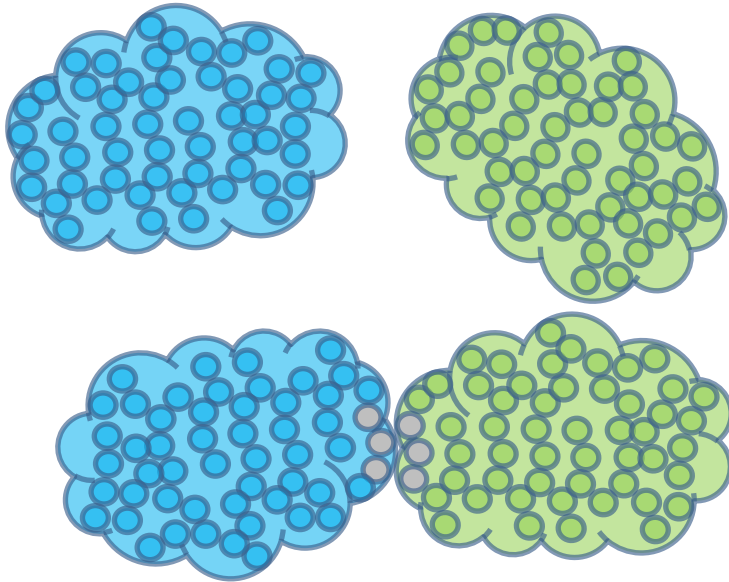


Medical Data Classification



Predicting Protein interactions, interfaces and affinity

- Input: Two protein structures or sequences
- Output: What residue pairs interact



PAIRpred: PARTner specific Interacting Residue PREDictor

What is PAIRpred?

We present a novel partner-specific protein-protein interaction site prediction method called PAIRpred. PAIRpred uses information from both proteins in a protein complex to predict pairs of interacting residues from the two proteins using both sequence and structure information. PAIRpred presents a more detailed model of protein binding, and offers state of the art accuracy in predicting binding sites at the protein level as well as inter-protein residue contacts at the complex level. PAIRpred is available as a cloud based webserver. You can use it to make predictions for proteins of your interest.

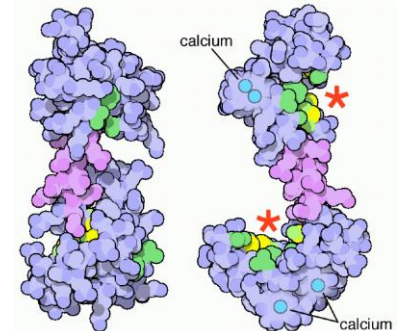
Apply PAIRpred to your proteins here

Input File Type:

File containing sequence or structure of ligand: No file chosen

File containing sequence or structure of receptor: No file chosen

Email:

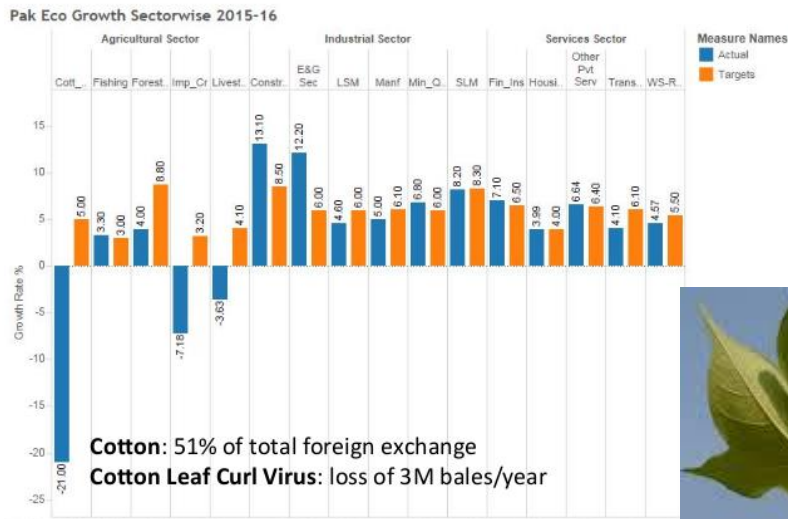
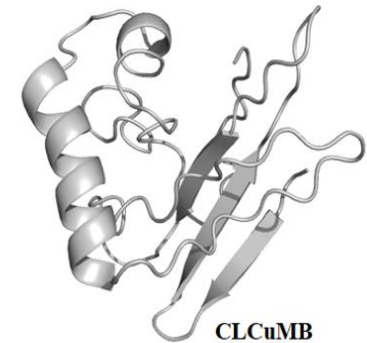
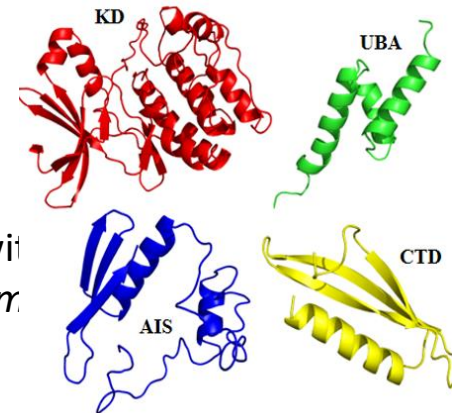


- [1] Fayyaz Minhas and Asa Ben-Hur, Multiple instance learning of Calmodulin binding sites. **Bioinformatics** 28, i416, 2012.
- [2] Wajid Abbasi, Amina Asif, Saiqa Andleeb, and Fayyaz Minhas, CaMELS: In silico Prediction of Calmodulin Binding Proteins and their Binding Sites, in **Proteins: Structure, Function and Bioinformatics**, 2017.
- [3] Issues In Performance Evaluation for Host-Pathogen Protein Interaction Prediction, Wajid A. Abbasi and Fayyaz Minhas, in **Journal of Bioinformatics and Computational Biology**, vol. 14, no. 3, 1650011, January 2016.
- [4] Training host-pathogen protein-protein interaction predictors , Abdul Hannan Basit, Wajid Arshad Abbasi, Amina Asif, Sadaf Gull and **Fayyaz Minhas**, in **Journal of Bioinformatics and Computational Biology**, Vol. 16, No. 04, 1850014 (2018).
- [5] Issues In Performance Evaluation for Host-Pathogen Protein Interaction Prediction, Wajid A. Abbasi and Fayyaz Minhas, in **Journal of Bioinformatics and Computational Biology**, vol. 14, no. 3, 1650011, January 2016.
- [6] Training host-pathogen protein-protein interaction predictors , Abdul Hannan Basit, Wajid Arshad Abbasi, Amina Asif, Sadaf Gull and **Fayyaz Minhas**, in **Journal of Bioinformatics and Computational Biology**, Vol. 16, No. 04, 1850014 (2018).

Experimental Validation: SnRK1- β C1

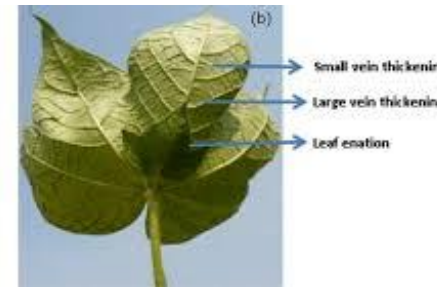
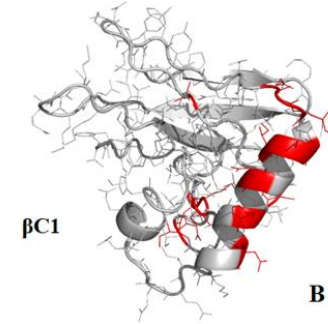
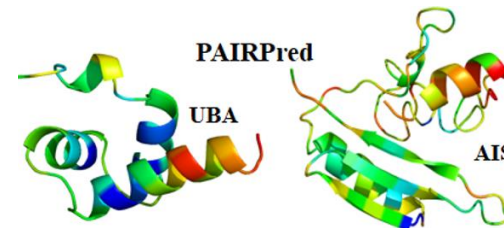
- In Silico Prediction and Validations of Domains Involved in *Gossypium hirsutum* SnRK1 Protein Interaction with Cotton leaf curl Multan betasatellite encoded β C1
- β C1, pathogenicity determinant encoded by Cotton leaf curl Multan betasatellite interacts with calmodulin-like protein 11 (CML11) in *Gossypium hirsutum*

GhSnRK1 506aa

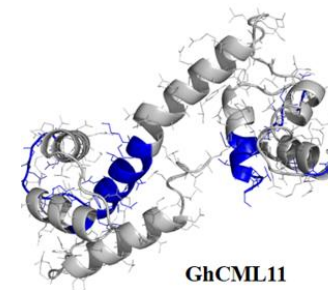


Cotton: 51% of total foreign exchange
Cotton Leaf Curl Virus: loss of 3M bales/year

Actual and Targets for each Sub-sec broken down by Sector. Color shows details about Actual and Targets. The marks are labeled by Actual and Targets.



4



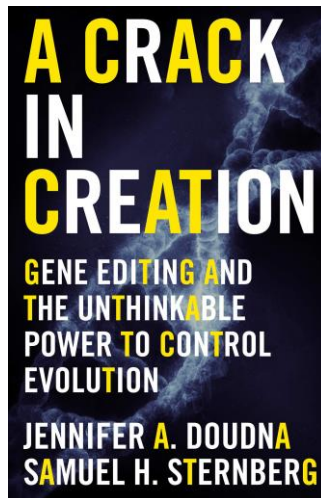
<http://www.dawn.com/news/1201323>

In Silico Prediction and Validations of Domains Involved in *Gossypium hirsutum* SnRK1 Protein Interaction with Cotton leaf curl Multan betasatellite encoded β C1, Kamal, Hira, Fayyaz ul Amir Afsar Minhas, Hanu Pappu, Imran Amin et al., in *Frontiers in Plant Science* 10 (2019): 656.

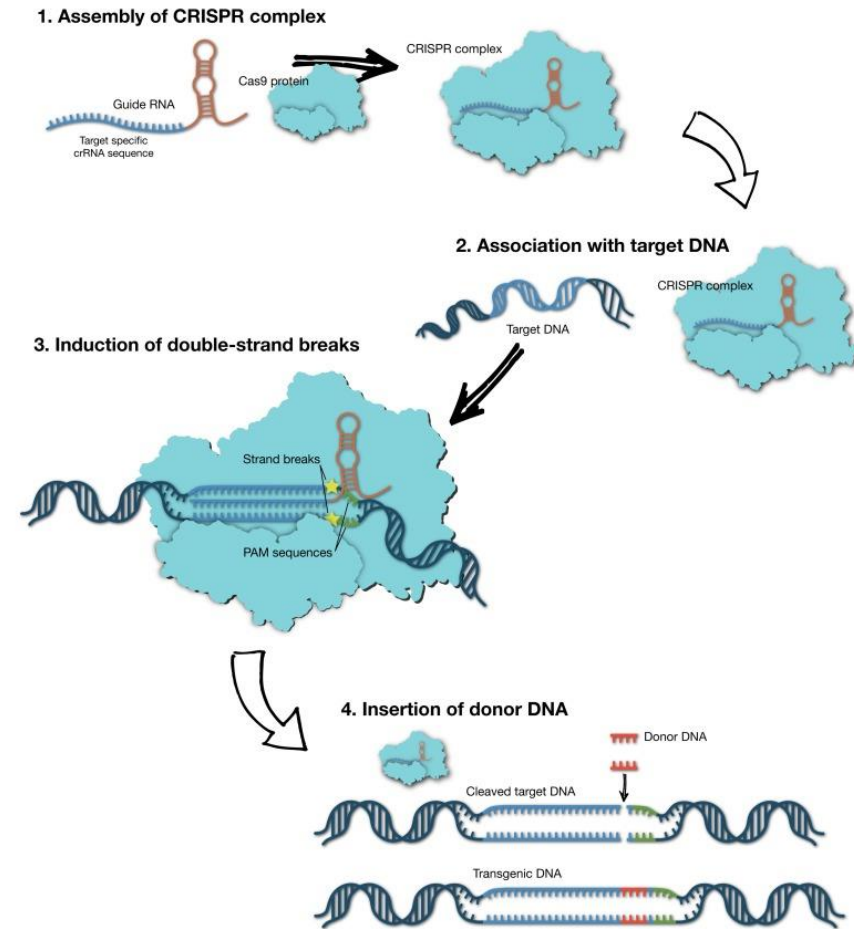
Bioinformatics and molecular analysis of *Gossypium hirsutum* calmodulin-like protein (CML11) interaction with begomovirus-transcription activator protein C2. Hira Kamal, Fayyaz Minhas, et al., in *PLoS One* (In press).

Predicting anti-CRISPR proteins

- 20 proteins for training
- Identified 3 new anti-CRISPR proteins using a ranking ML model

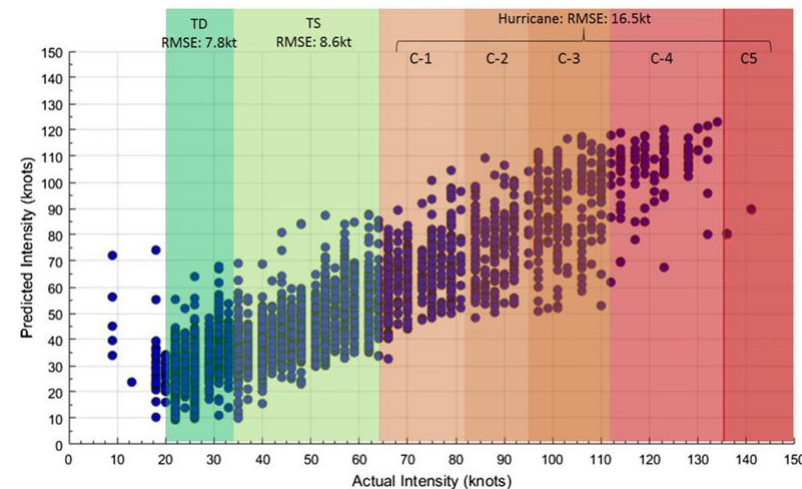
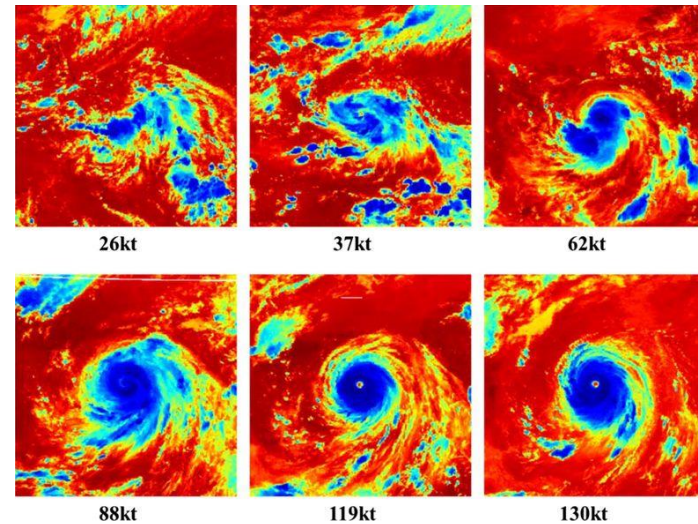


Jennifer Doudna



Data Science: Hurricane Intensity Prediction

- Hurricane Intensity Prediction
 - In collaboration with National Hurricane Center, USA
 - Deep-PHURIE



Deep-PHURIE: Deep Learning based Hurricane Intensity Estimation from Infrared Satellite Imagery, M. Dawood, A. Asif and Fayyaz Minhas, in Neural Computing and Applications.

pp. DOI: [10.1007/s00521-019-04410-7](https://doi.org/10.1007/s00521-019-04410-7), July 2019.

PHURIE: Hurricane Intensity Estimation from Infrared Satellite Imagery using Machine Learning, Amina Asif, Muhammad Dawood, Bismillah Jan, Javaid Khurshid, Mark DeMaria, and Fayyaz ul Amir Afsar Minhas, in Neural Computing and Applications, DOI: [http://dx.doi.org/10.1007/s00521-018-3874-6](https://doi.org/10.1007/s00521-018-3874-6), 2018

Data Science: Journal Recommendation System



PRISM: Journal Paper Recommendation System

PRISM helps you find journals that could be best suited for publishing your research papers. Please also consult the journal's Aims and Scope for further guidance. Ultimately, the Editor will decide on how well your article matches the journal. Simply insert your title and abstract and select the appropriate field-of-research for the best results.

Hover-Net: Simultaneous segmentation and classification c ✓

Nuclear segmentation and classification within Haematoxylin & Eosin stained histology images is a fundamental prerequisite in the digital pathology work-flow. The development of automated methods for nuclear segmentation and classification enables the quantitative analysis of tens of thousands of nuclei within a whole-slide pathology image, opening up possibilities of further analysis of large-scale nuclear morphometry. However, automated nuclear segmentation and classification is faced with a major

Clear

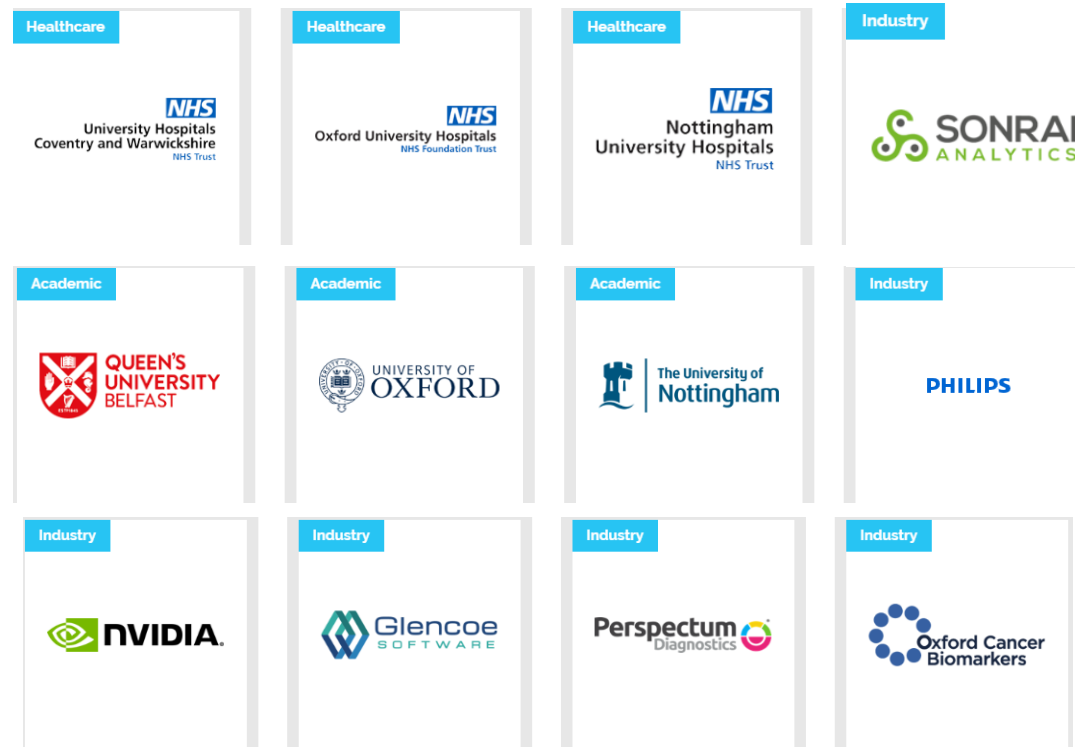
Find Journal

Our Recommendations
medical image analysis
ieee transactions on medical imaging
medical image computing and computer assisted intervention – miccai 2018
ieee transactions on geoscience and remote sensing
ieee transactions on image processing

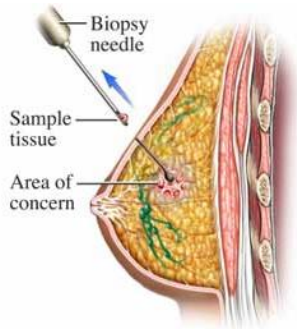
- Using classical NLP
- Using BERT

Current Focus: PATHLake

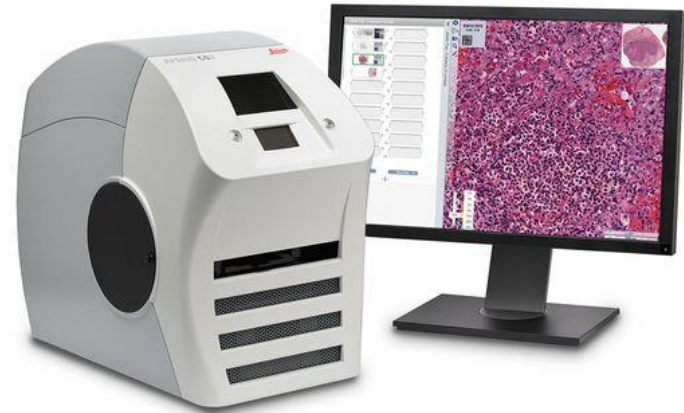
- PATHology data Lake, Analytics, Knowledge and Education
- UK Research and Innovation
- £15.7 million
- Objective:
 - Improve speed and accuracy of cancer diagnosis



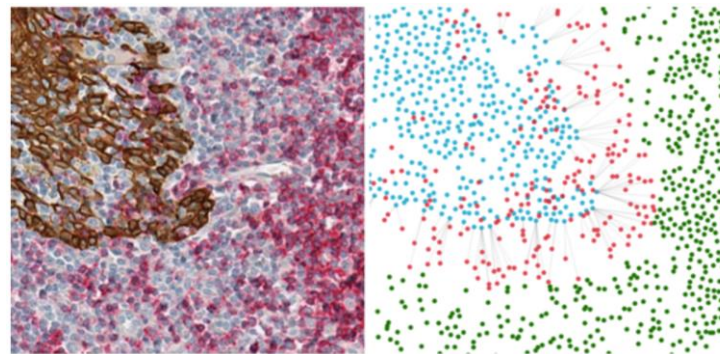
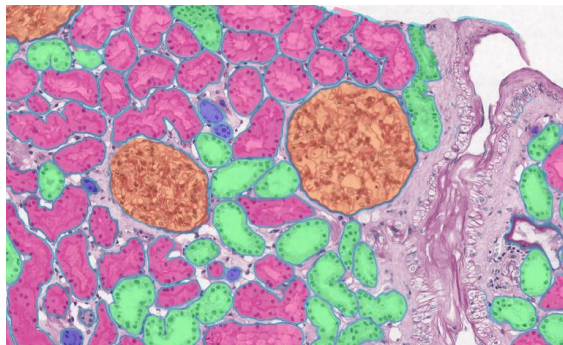
The Revolution in Pathology



Conventional Microscope Pathology



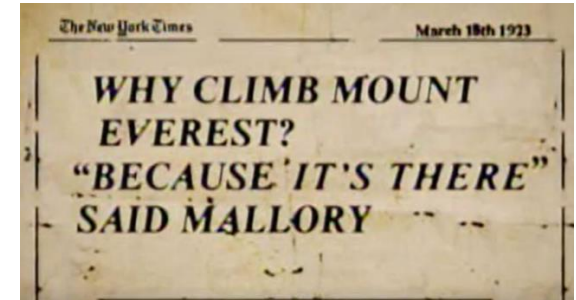
Digital Pathology



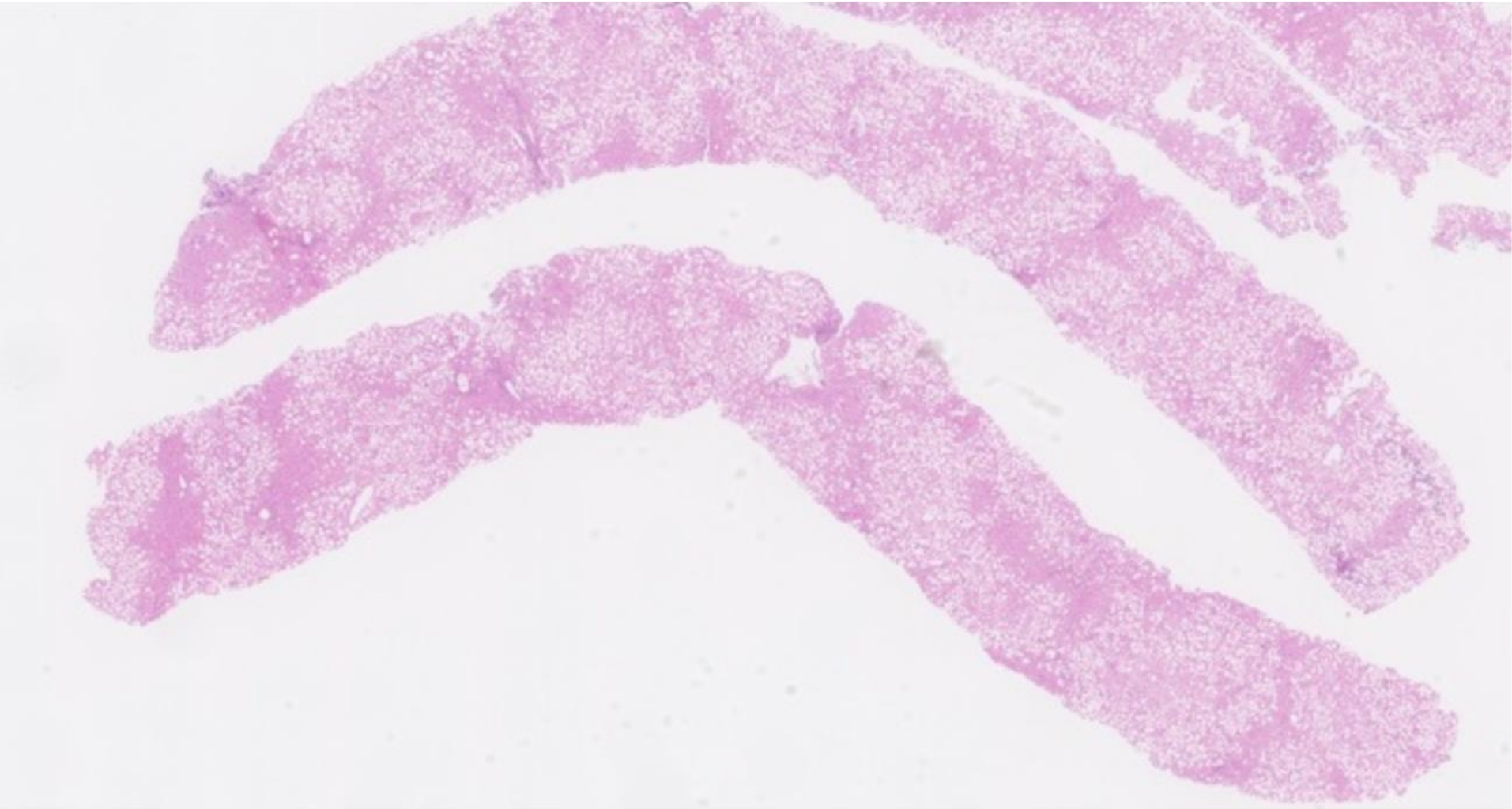
Computational Pathology

Why?

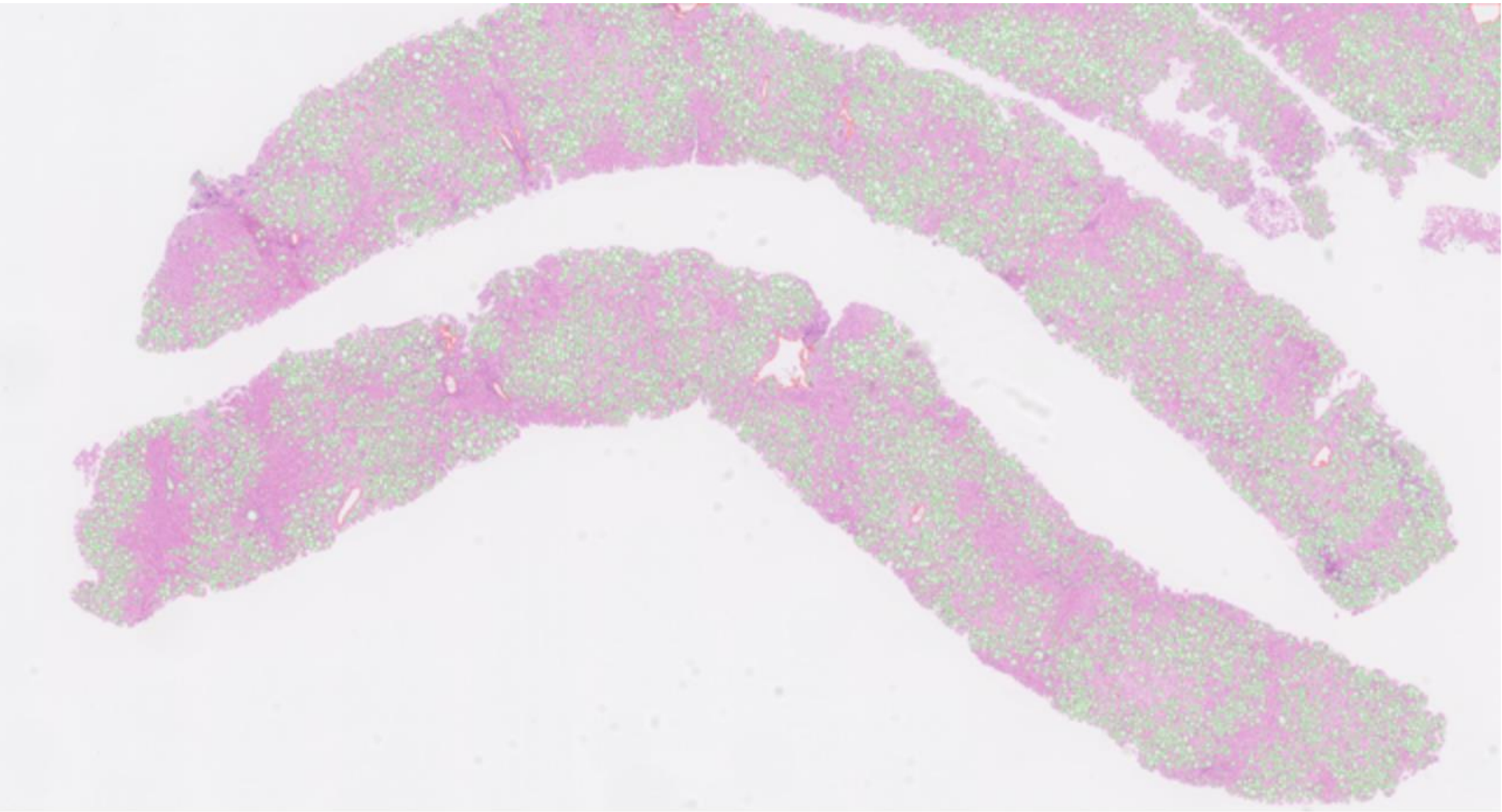
- Shortage of Pathologists
- Quantification is difficult
- Subjectivity
- Inter-observer variability



How much fat?

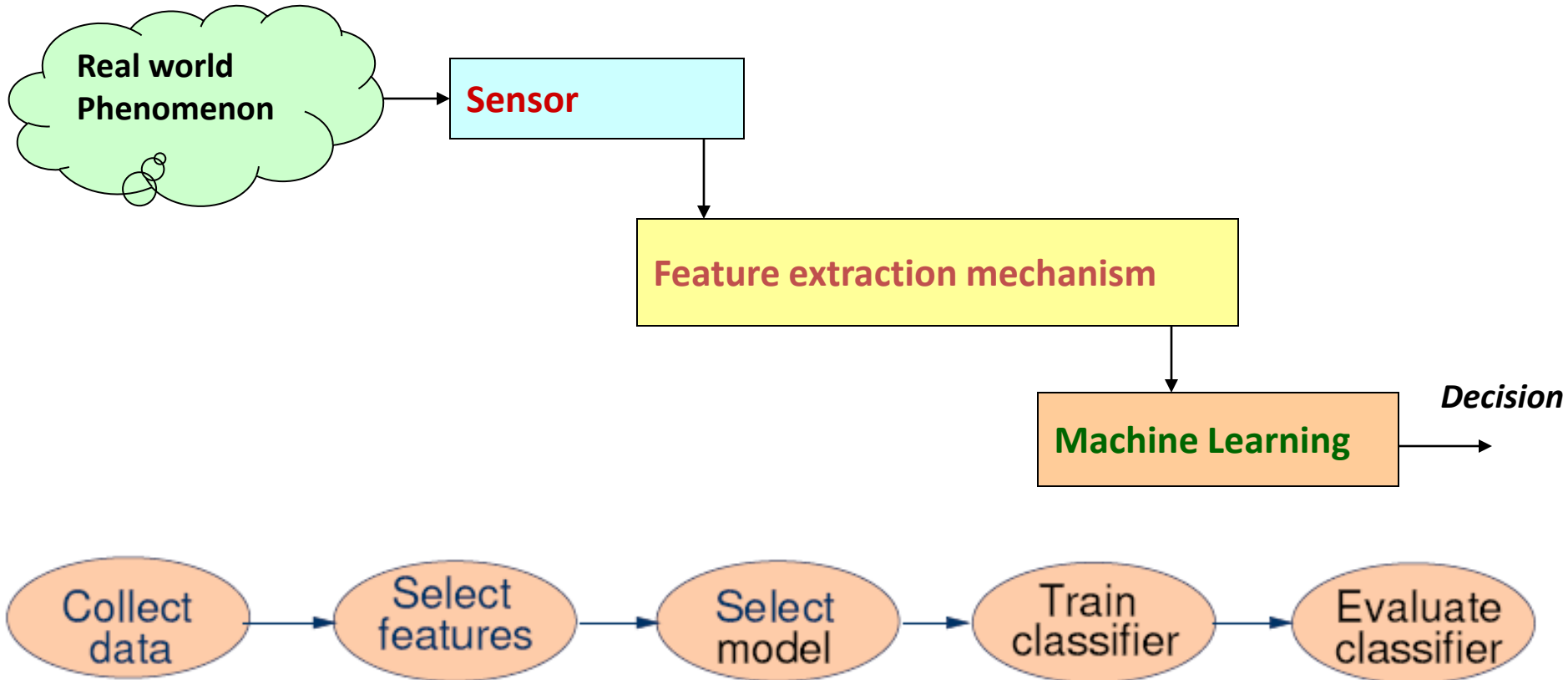


45.9%



Constructs of a PR System

- Identify the objective
 - Identify the unit of classification (example)
 - Image block, protein sequence,



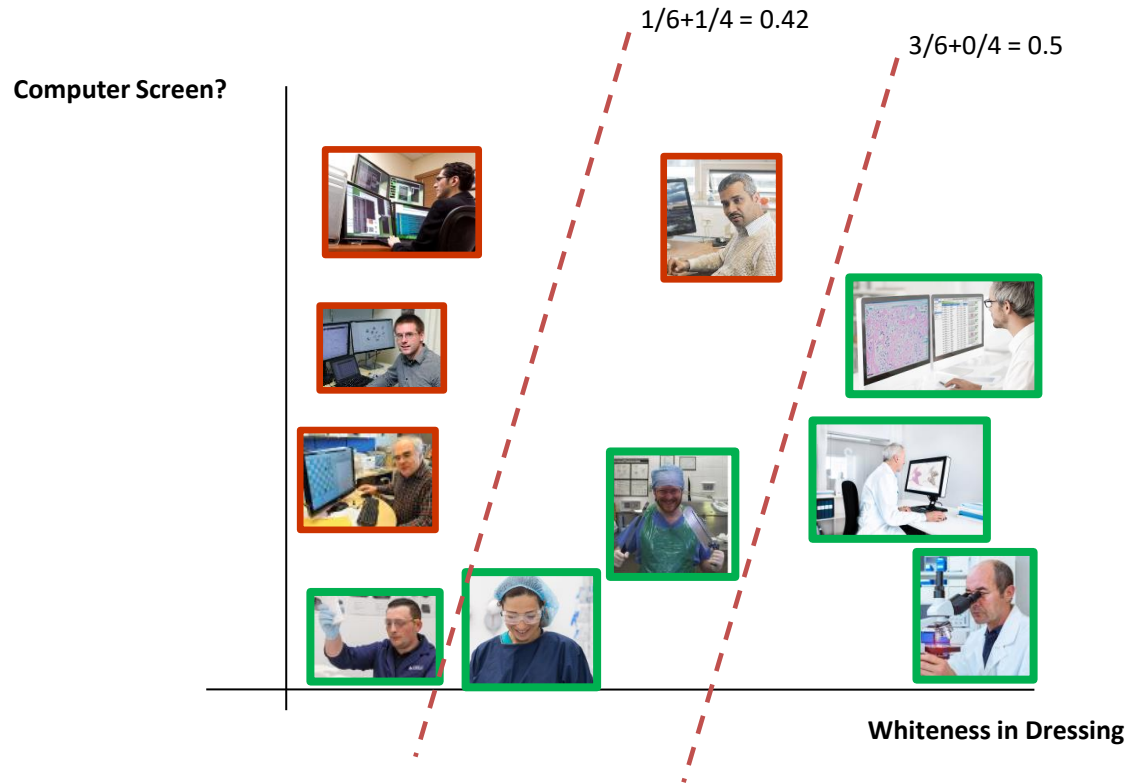
Learning from Data

- Example Case
 - Pathologists vs. Computer Scientists
 - Hypothetical!
 - Classify a person in their "native" environment

Constructs

- Sensor(s)
 - Camera
- Feature Extraction
 - White coats?
 - Computer Screen?
 - Income?
- Machine Learning

Feature Space



Feature Space Classification

Computer Screen?

$$0/6 + 0/4 = 0.0$$

