Explainable AI for the (Not-Always-Expert) Clinical Researcher

John H. Holmes, PhD
University of Pennsylvania Perelman School of Medicine
Department of Biostatistics, Epidemiology, and Informatics
Institute for Biomedical Informatics
Center for Clinical Epidemiology and Biostatistics

Perelman School of Medicine University of Pennsylvania

DBEI Department of Biostatistics Epidemiology & Informatics

CCEB

Institute for Biomedical Informatics

Outline

- Setting the stage: The Premise
- A little AI history (and who doesn’t like that?)
- Proposal of a data analytics framework (notice the “a”…. It’s not “the”)
- Instantiating the framework: PennAI (what we’ve been up to)
- Summary

How does AI serve clinical research?

- The challenge
  - Clinical data are complex, heterogeneous, and difficult to understand
  - Clinical research usually framed in the context of hypotheses
  - Clinical data often need enrichment and novel data science and analytic approaches to answer research questions

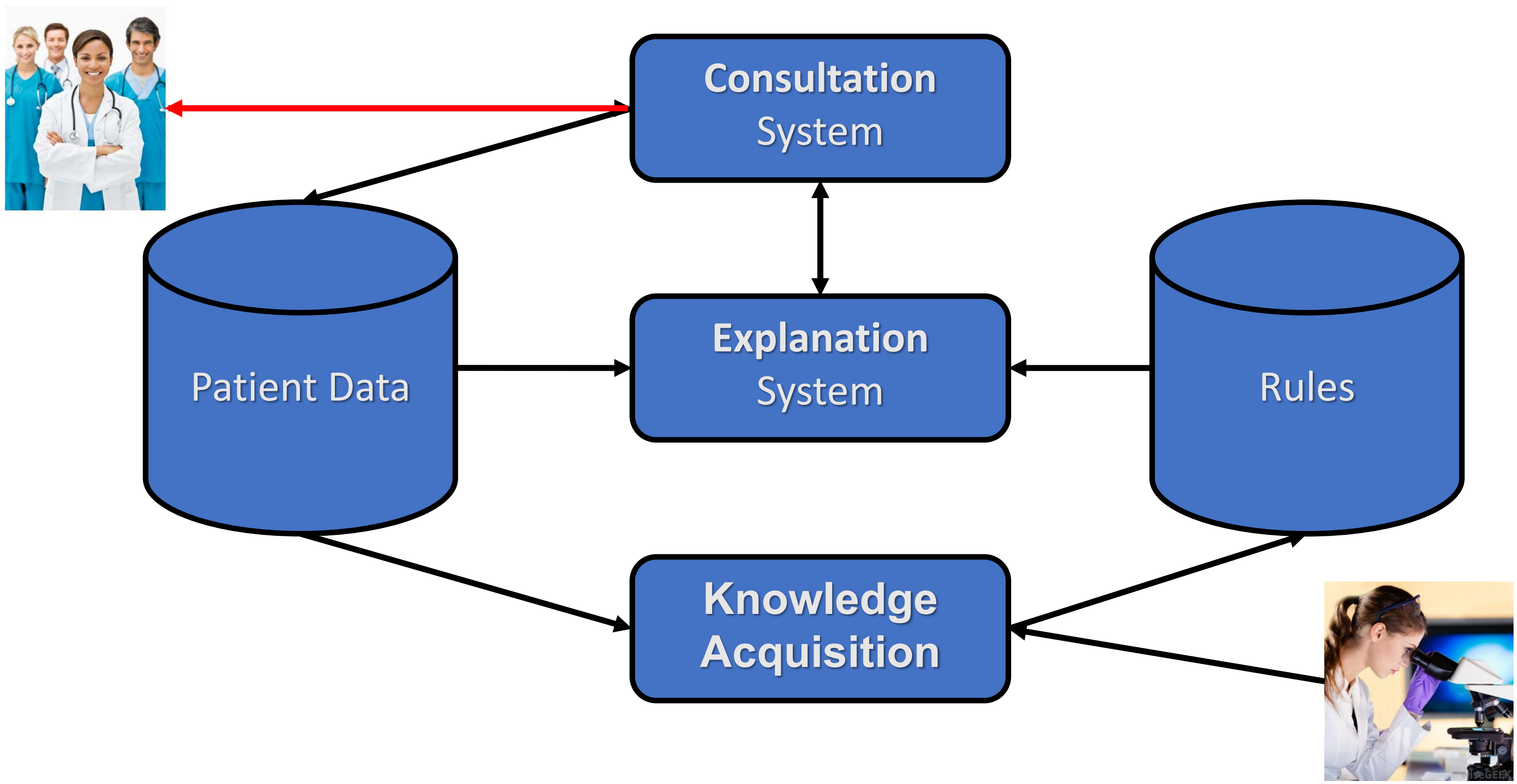
- The opportunity
  - To develop specialized tools for clinical research analytics informed by best practices and novel methods in artificial intelligence
  - To provide the clinical researcher with access to these tools
  - To identify disease entities, their characteristics, and potential causal pathways
  - To propose and evaluate therapies
  - **To make AI methods and results *explainable* to the researche**r!

The Premise: AI should be open, easy, and accessible

Where AI started: Mimicking human cognition

- Rules (expert systems)
- Frames (case-based reasoning)
- Cognitive computing (default hierarchies)
- All focused on capturing human knowledge and reasoning in some type of static representation that may be updated (or not) with experience

The first medical expert system: MYCIN



A closer look at explanation in MYCIN

- Displays the rule being invoked at any point during the consultation or inferential chain
- Keeps a history of rule invocation and associates each rule with questions asked of user
- Uses a “goal tree” to illustrate inference

An example of the MYCIN explanation facility

32) Was penicillinase added to this blood culture
(CULTURE-1)?
**WHY
[i.e. WHY is it important to determine whether penicillinase was added to CULTURE-1?]

> [3.0] This will aid in determining whether ORGANISM-1 is a contaminant. It has already been established that
> [3.1] the site of CULTURE-1 is blood, and
> [3.2] the gram stain of ORGANISM-1 is grampos
> Therefore, if
>
> > [3.3] penicillinase was added to this blood culture then there is weakly suggestive evidence that…

https://www.slideshare.net/vini89/mycin?qid=187f4b1b-e711-4a36-943b-05e01f02ad0b&v=&b=&from_search=3

Did MYCIN work?

Pretty well, at least for validation/verification

Ratings of Antimicrobial Selection Based on Evaluator Rating and Etiologic Diagnosis

Prescribers	No. (%) of Cases in Which Therapy Rated Acceptable* by an Evaluator (n=80)	No. (%) of Cases in Which Therapy Rated Acceptable* by Majority of Evaluators (n=10)	No. of Cases in Which Therapy Failed to Cover a Treatable Pathogen (n=10)
MYCIN	52 (65)	7 (70)	0
Faculty, 1	50 (62.5)	5 (50)	1
Faculty, 2	48 (60)	5 (50)	1
Infectious diseases fellow	48 (60)	5 (50)	1
Faculty, 3	46 (57.5)	4 (40)	0
Actual therapy	46 (57.5)	7 (70)	0
Faculty, 4	44 (55)	5 (50)	0
Resident	36 (45)	3 (30)	1
Faculty, 5	34 (42.5)	3 (30)	0
Student	24 (30)	1 (10)	3

Yu VL, et al.: Antimicrobial selection by a computer: A blinded evaluation by infectious disease experts. *JAMA* 242(12):1279-82 (1979)

However, MYCIN was never used in practice…

- Required extensive knowledge engineering and rules updating
- Legal and regulatory issues (that persist to this day)
- Even with an explanation facility, it was cumbersome to use

Another type of explanation: Multivariable modeling

$$q = 0.0112 \times 1.059^{(\text{age at diagnosis} - 55)} \times 0.525^{\text{female}}$$
$$\times 1.350^{\text{smoker}} \times 1.144^{(\text{HbA}_{1c} - 6.72)}$$
$$\times 1.073^{((\text{bpsys} - 135.7)/10)}$$
$$\times 3.1105^{(\log(\text{chol} \div \text{HDL}) - 1.59)}$$

$$\text{UKPDS 10-year CHD risk}\,(\%)$$
$$= \left(1 - \exp\left(-q \times \left(\left(1 - 1.0785^{10}\right) / \left(1 - 1.078\right)\right)\right)\right) \times 100$$

Metcalf PA, Wells S, Jackson RT: Assessing 10-year coronary heart disease risk in people with Type 2 diabetes mellitus: Framingham versus United Kingdom Prospective Diabetes Study. *J Diab Mellitus* 4:1 (2014)
https://www.scirp.org/journal/PaperInformation.aspx?PaperID=41893

Explanation has eluded us for many machine learning approaches

Random forests

Neural networks

Support vector machines

...

A serious question:
Is explainability a *desideratum* or a *sine qua non* in biomedical applications of AI?

Darlington KW: Designing for Explanation in Health Care Applications of Expert Systems. *Sage Open* 2011, DOI: 10.1177/2158244011408618. https://journals.sagepub.com/doi/pdf/10.1177/2158244011408618

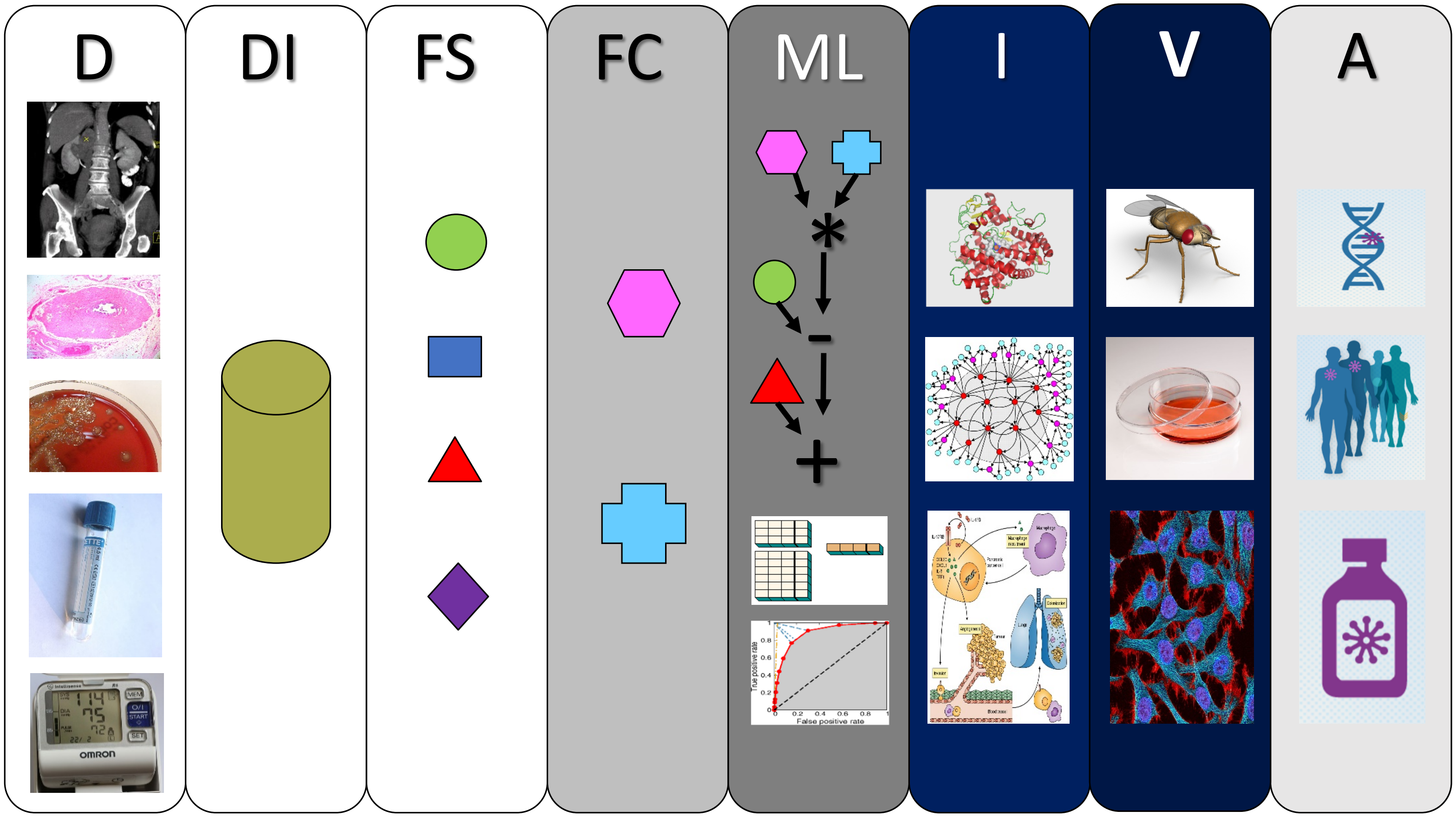
What we really could use is…

A conceptual framework for integrating AI in biomedical data analytics that is sensitive to users' needs and expertise

(… in other words, it makes room for explainability)

With thanks to Jason H. Moore, PhD

The framework: A Data analytics pipeline

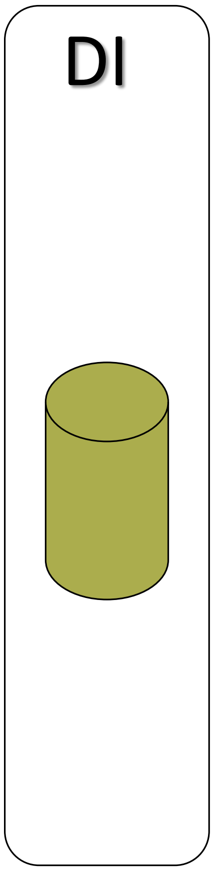


Big Data

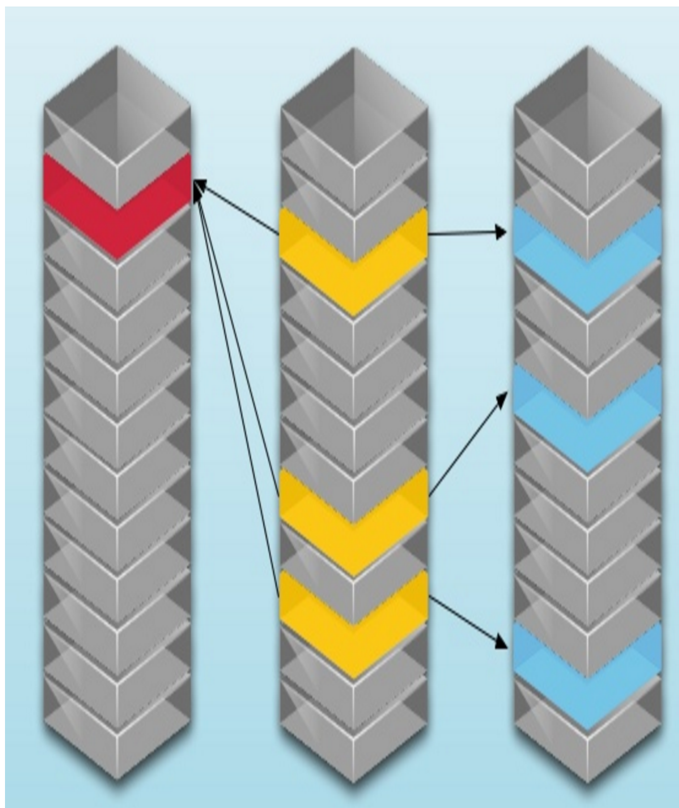
D

http://www.kdnuggets.com/

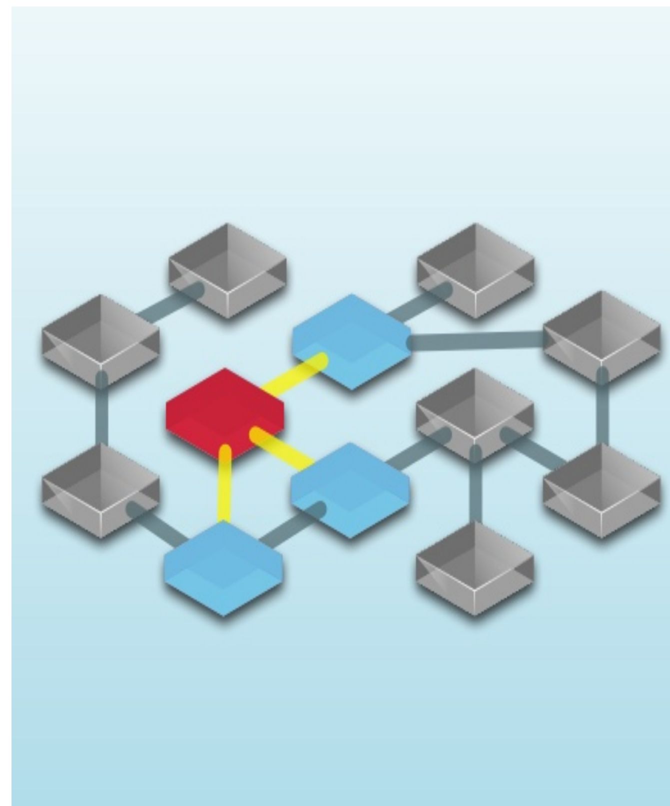
Data integration



Relational Database

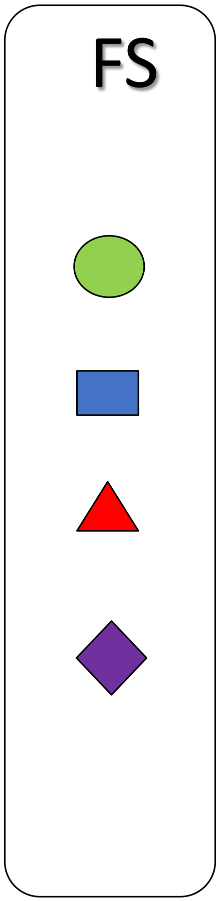


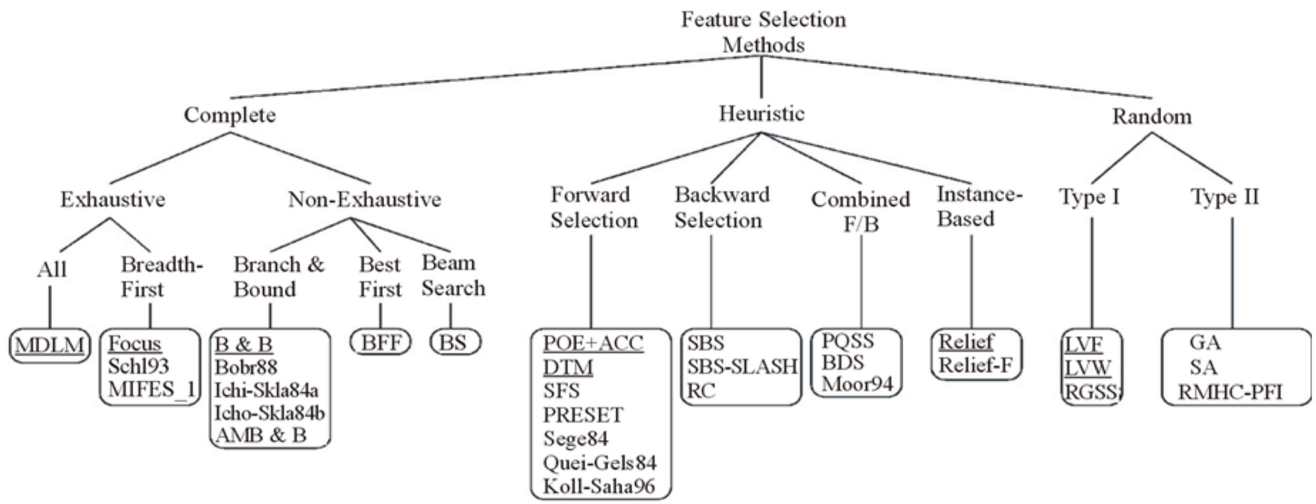
Graph Database



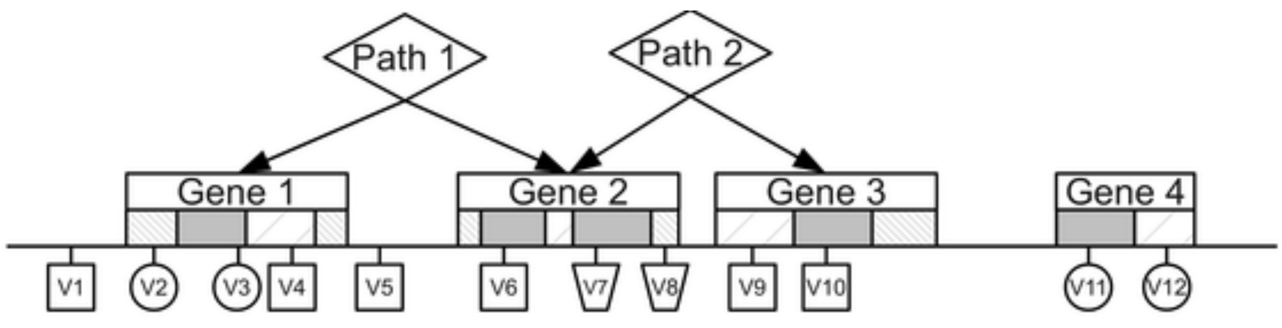
Michael Hunger – Neo4j

Feature selection



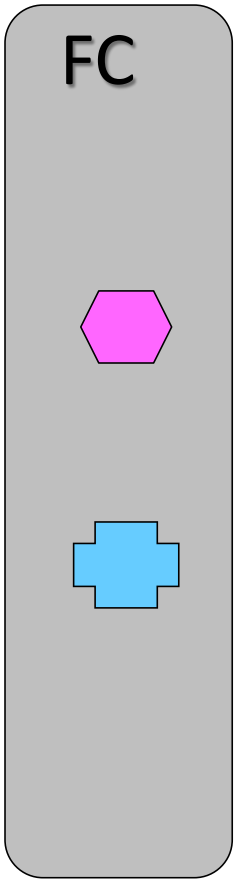


Sohangir – J Soft Engin App (2013)



Ritchie – PLoS Genetics (2013)

Feature construction



		X_1: 0	1	2
X_2	0	1	2	3
	1	4	5	6
	2	7	8	9

→ Z_1

1	2	3	4	5	6	7	8	9

		X_3: 0	1	2
X_4	0	0	0	0
	1	0	1	1
	2	0	1	1

→ Z_2

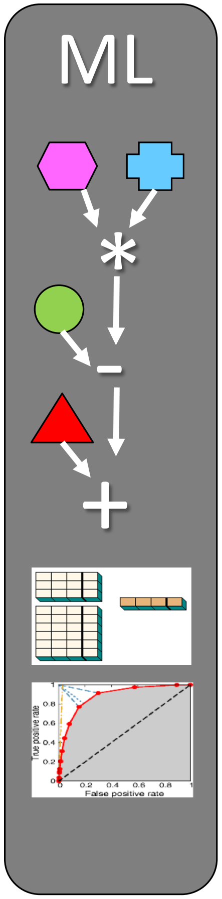
0	1

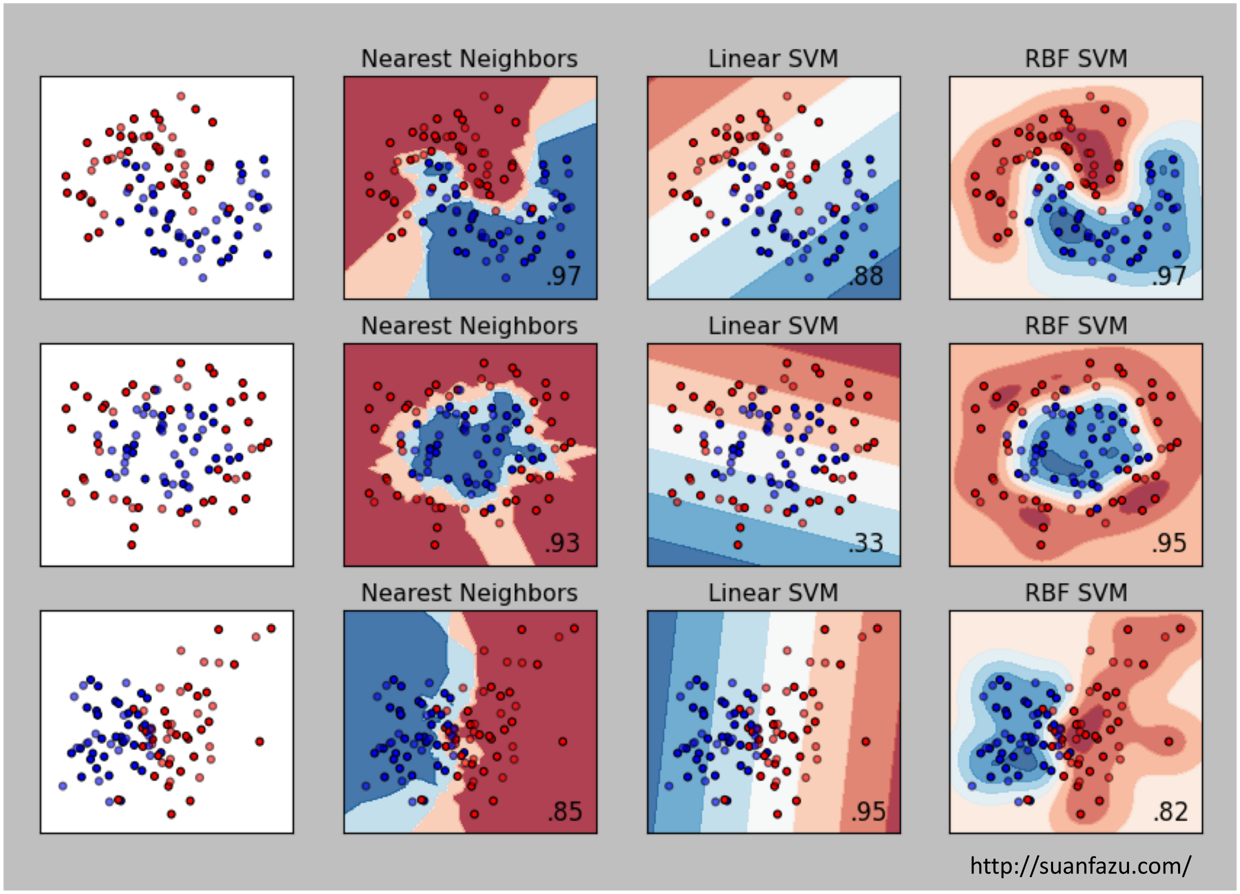
		X_5: 0	1	2
X_6	0	1	0	1
	1	0	1	0
	2	1	0	1

→ Z_3

0	1

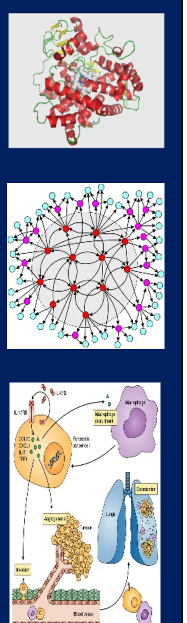
Machine learning

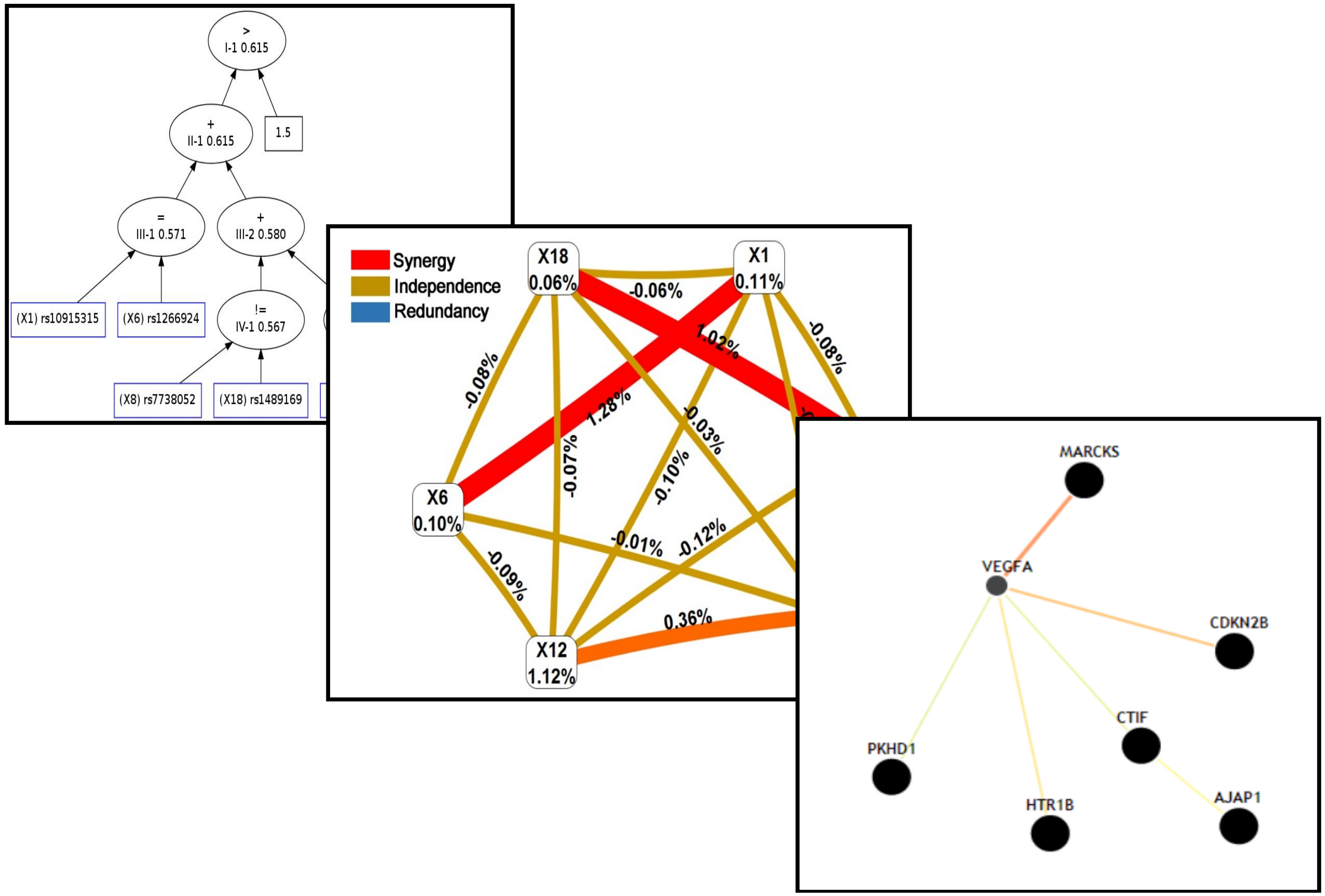




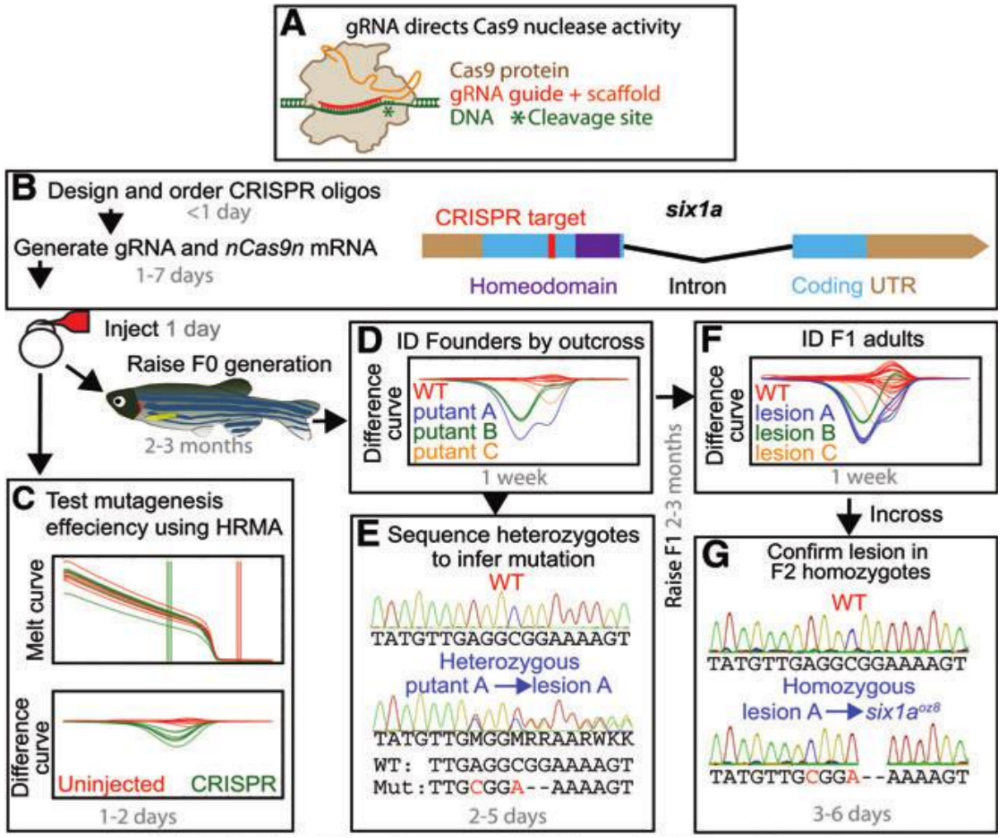
Statistical and biological interpretation

I

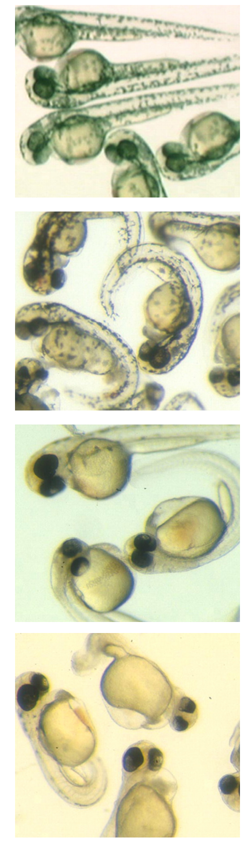




Biological validation

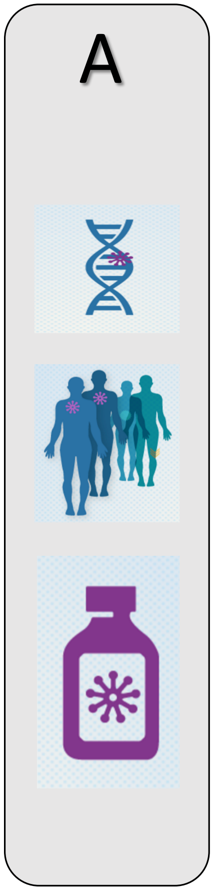


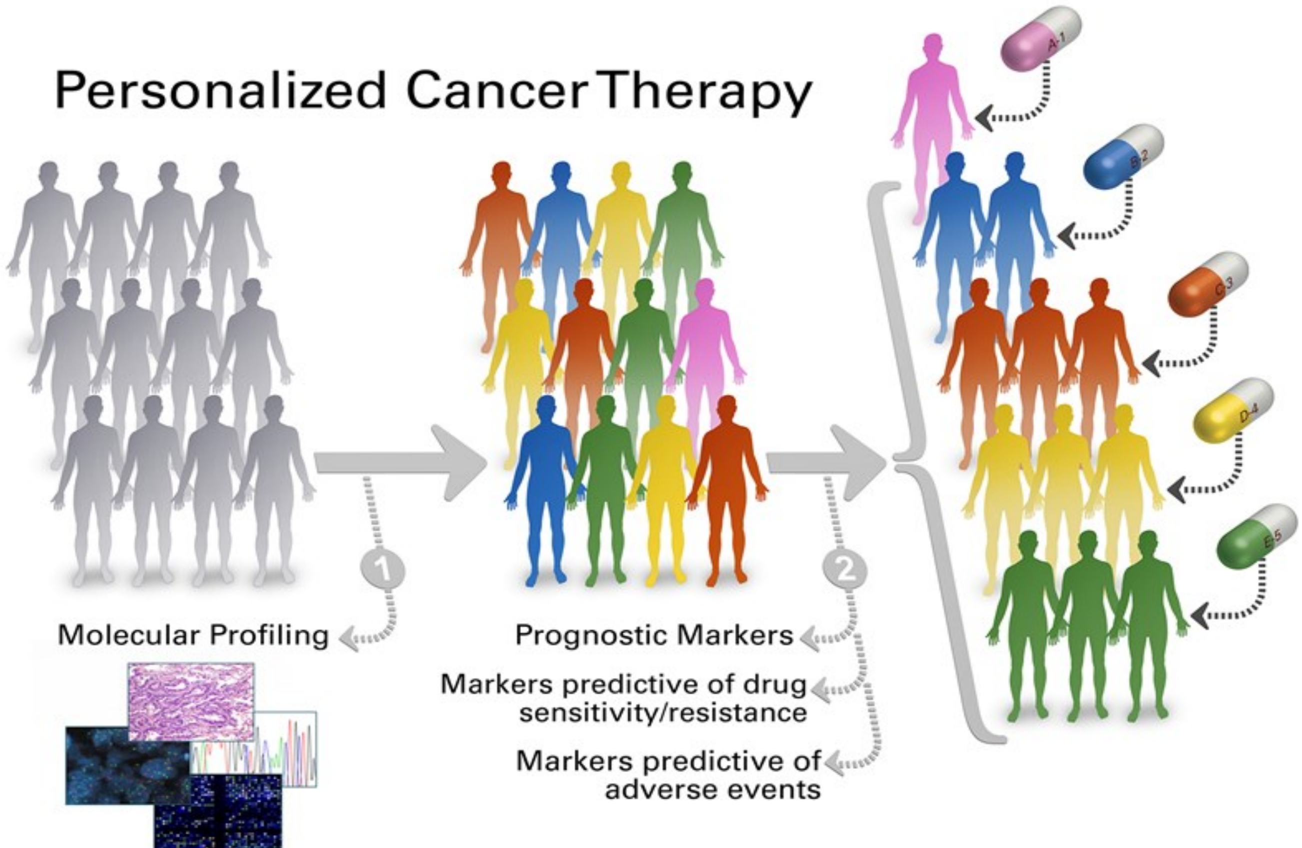
Talbot, Zebrafish (2014)



dev.biologists.org

Clinical application



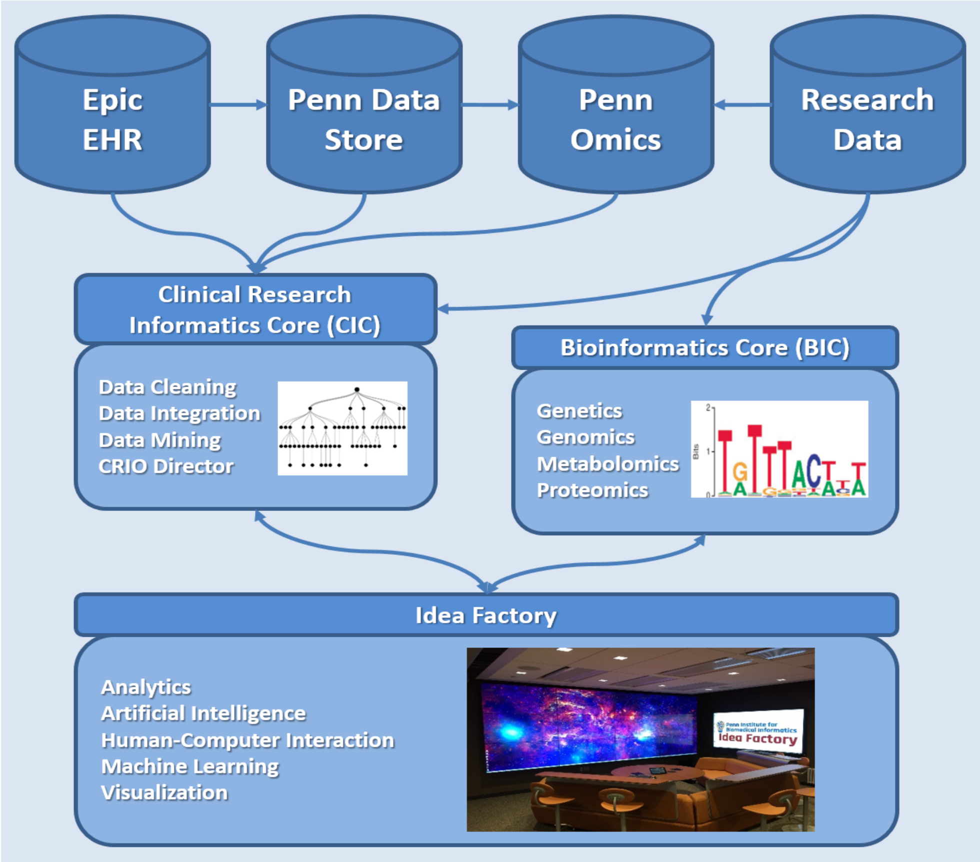


M.D. Anderson

What we are doing at the University of Pennsylvania: PennAI

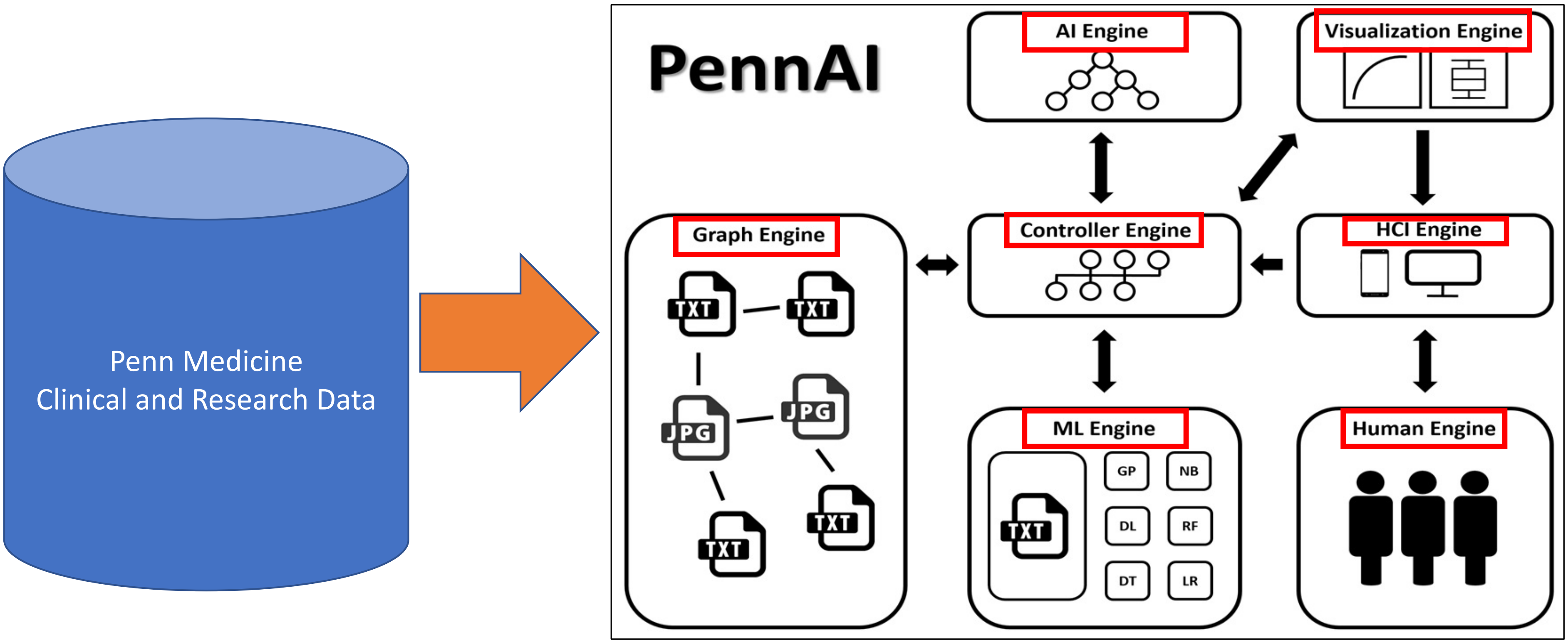
AI-assisted machine learning-driven hypothesis and knowledge discovery for clinical research

Penn's Informatics Infrastructure Overview



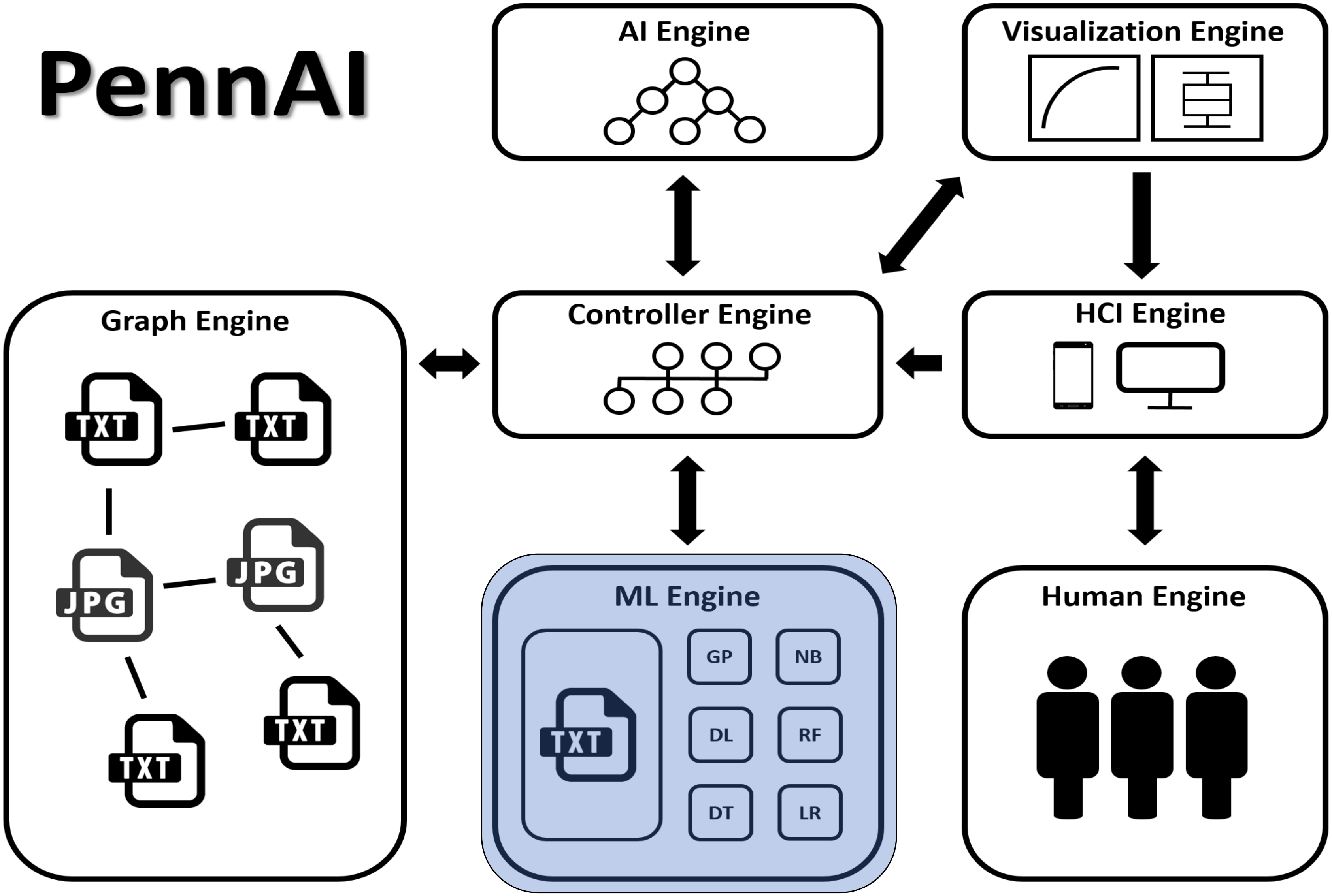
Penn IBI Idea Factory

Our machine learning environment: PennAI



www.pennai.org

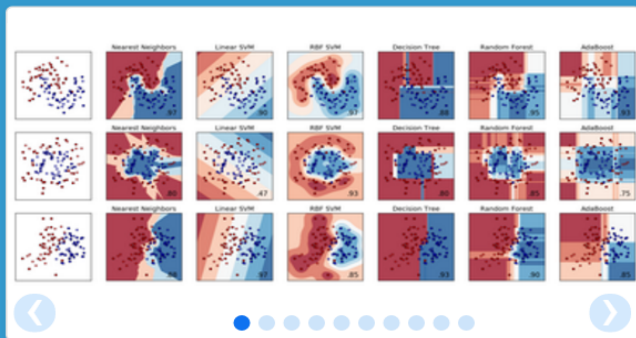
PennAI



AI Engine
Visualization Engine
Graph Engine
Controller Engine
HCI Engine
TXT
TXT
JPG
JPG
TXT
TXT
ML Engine
TXT
GP
NB
DL
RF
DT
LR
Human Engine

Home Installation Documentation Examples

http://scikit-learn.org/



scikit-learn

Machine Learning in Python

- Simple and efficient tools for data mining and data analysis
- Accessible to everybody, and reusable in various contexts
- Built on NumPy, SciPy, and matplotlib
- Open source, commercially usable - BSD license

Classification

Identifying to which category an object belongs to.

Applications: Spam detection, Image recognition.
Algorithms: SVM, nearest neighbors, random forest, ... — Examples

Regression

Predicting a continuous-valued attribute associated with an object.

Applications: Drug response, Stock prices.
Algorithms: SVR, ridge regression, Lasso, ... — Examples

Clustering

Automatic grouping of similar objects into sets.

Applications: Customer segmentation, Grouping experiment outcomes
Algorithms: k-Means, spectral clustering, mean-shift, ... — Examples

Dimensionality reduction

Reducing the number of random variables to consider.

Applications: Visualization, Increased efficiency
Algorithms: PCA, feature selection, non-negative matrix factorization. — Examples

Model selection

Comparing, validating and choosing parameters and models.

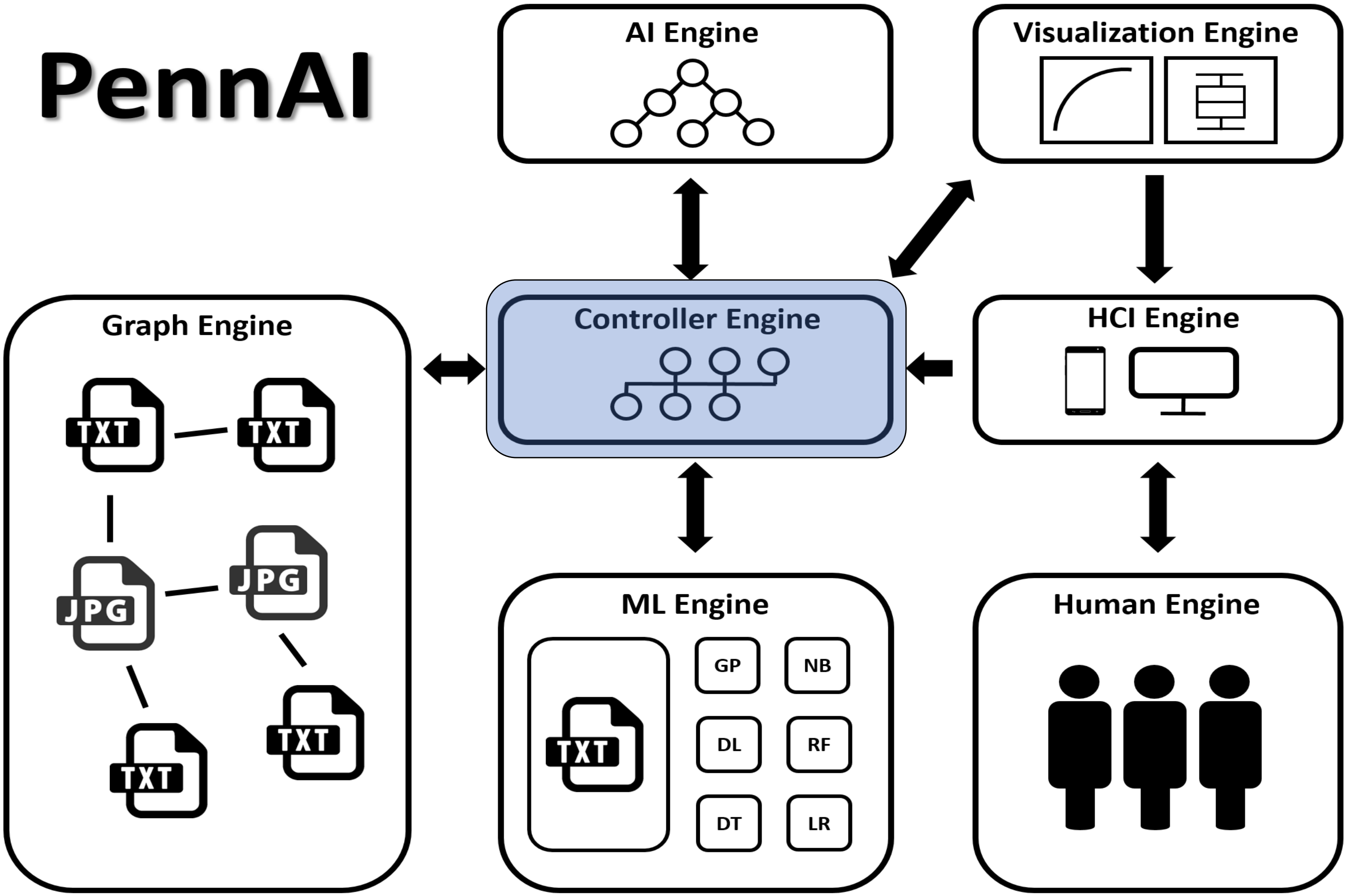
Goal: Improved accuracy via parameter tuning
Modules: grid search, cross validation, metrics. — Examples

Preprocessing

Feature extraction and normalization.

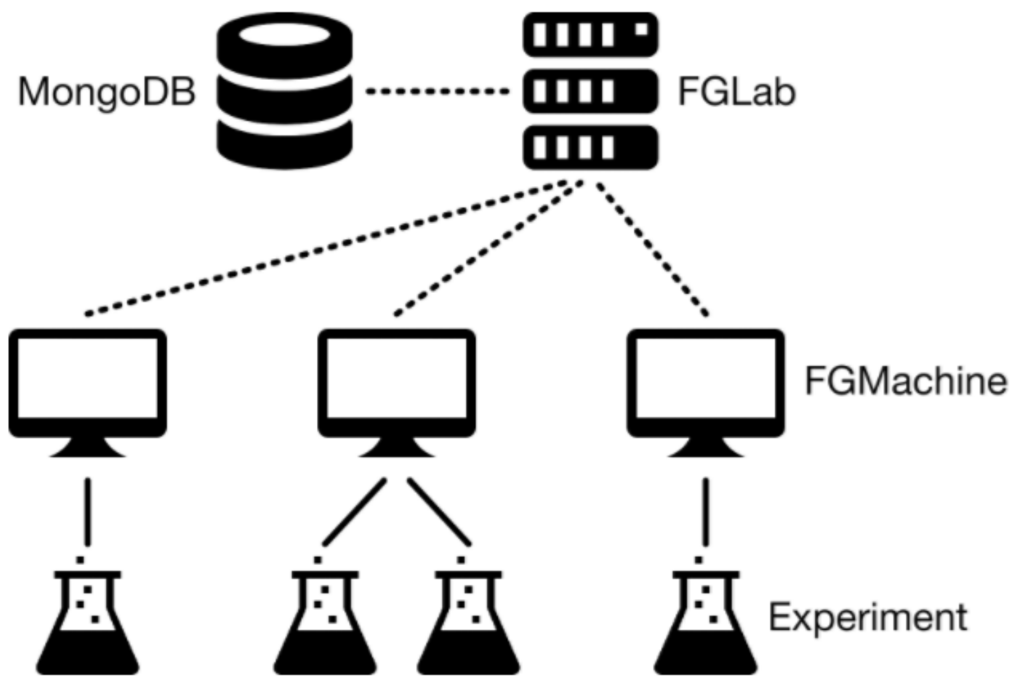
Application: Transforming input data such as text for use with machine learning algorithms.
Modules: preprocessing, feature extraction. — Examples

PennAI



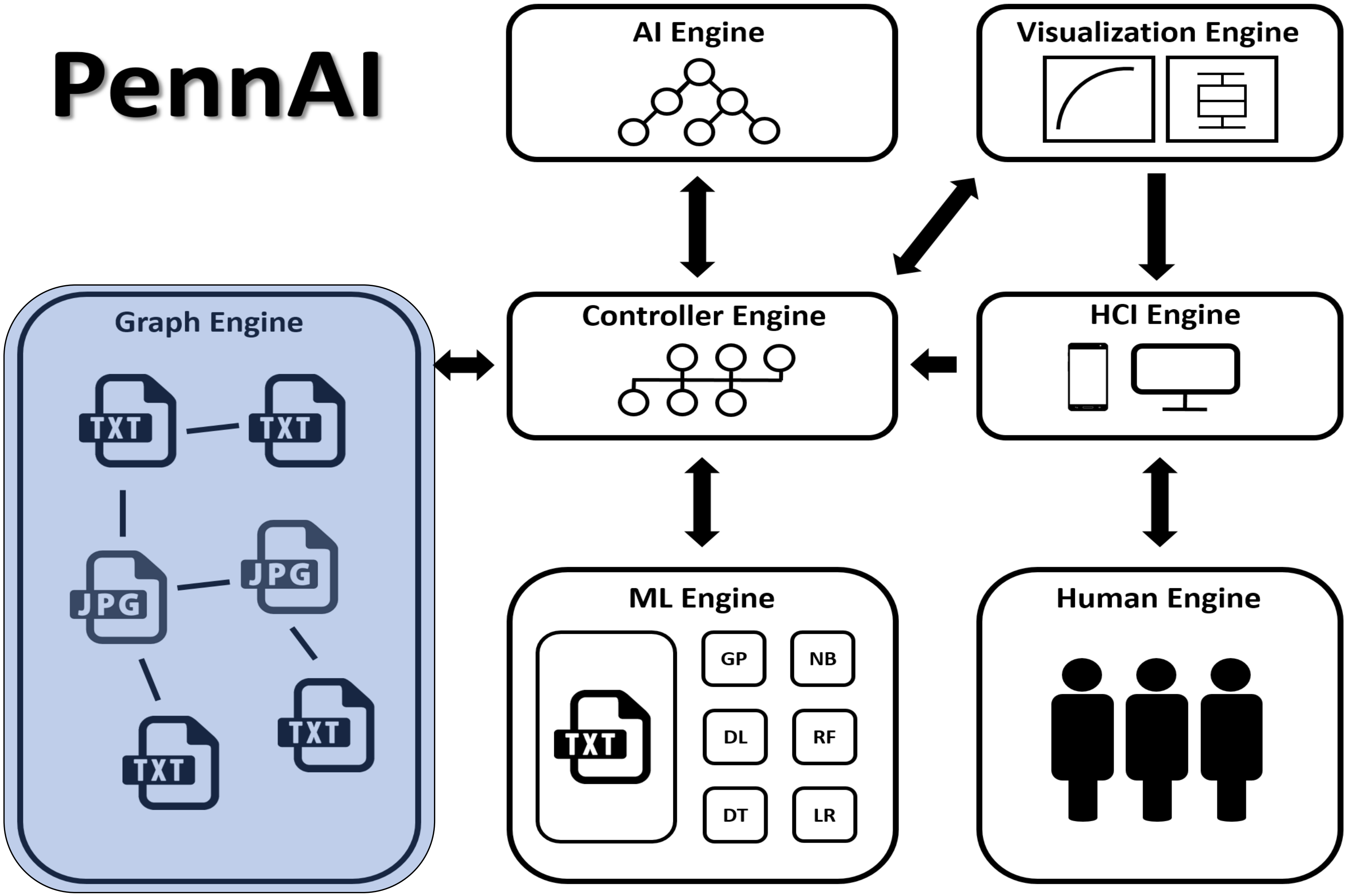
AI Engine
Visualization Engine
Graph Engine
TXT
TXT
JPG
JPG
TXT
TXT
Controller Engine
HCI Engine
ML Engine
TXT
GP
NB
DL
RF
DT
LR
Human Engine

Controller: Future Gadget Lab

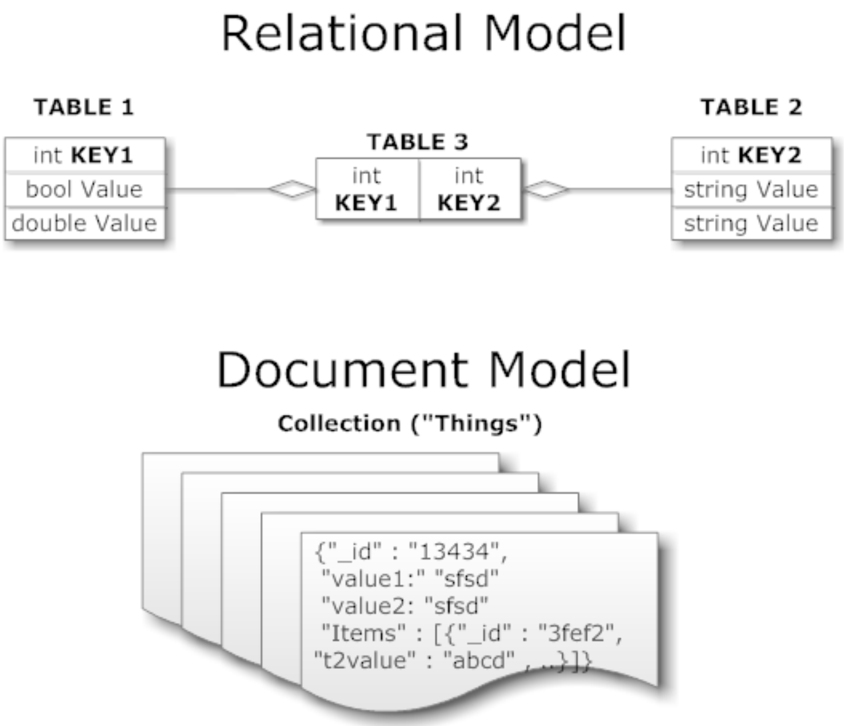


https://kaixhin.github.io/FGLab/

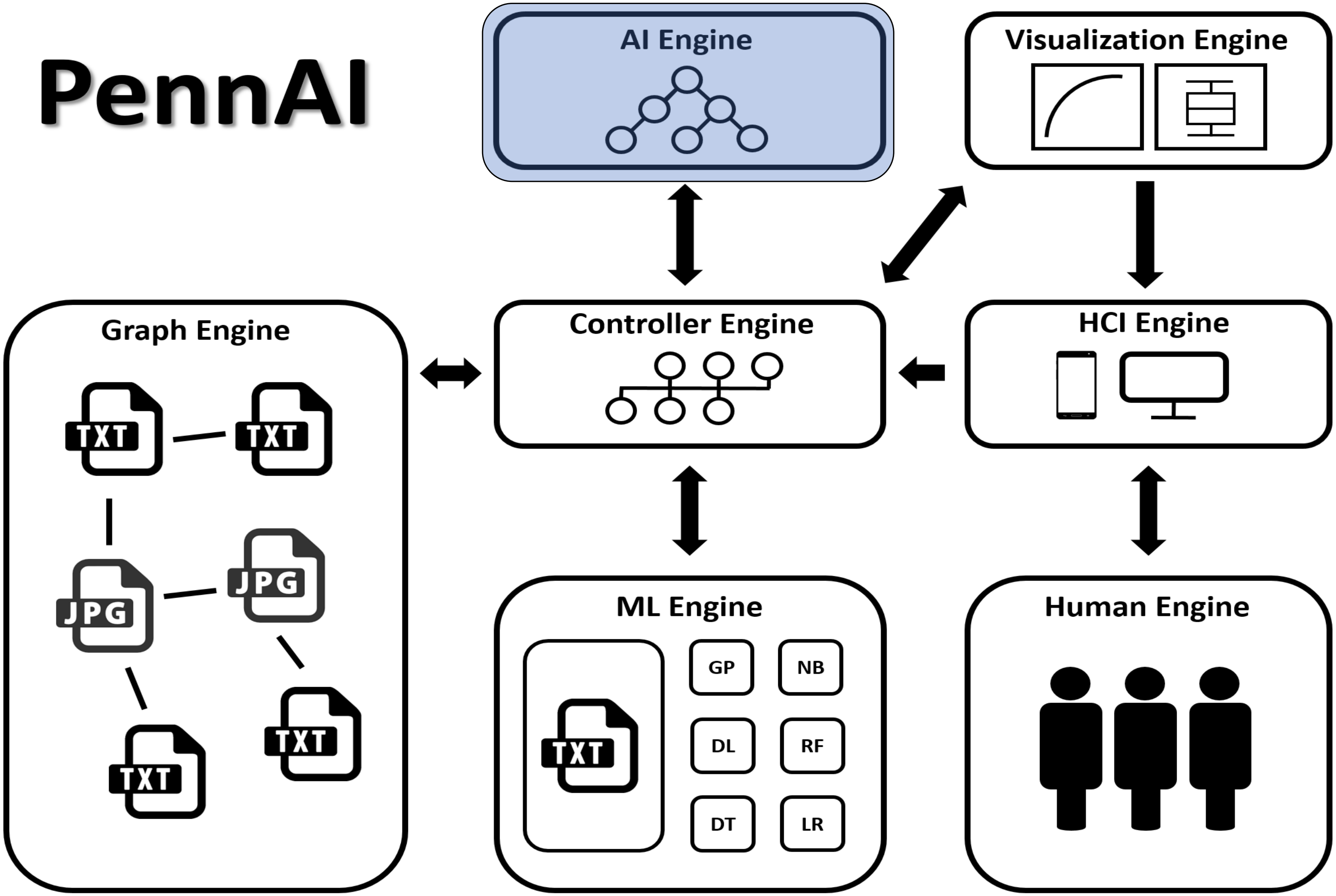
PennAI


AI Engine
Visualization Engine
Graph Engine
TXT
TXT
JPG
JPG
TXT
TXT
Controller Engine
HCI Engine
ML Engine
TXT
GP
NB
DL
RF
DT
LR
Human Engine

Database: MongoDB



PennAI

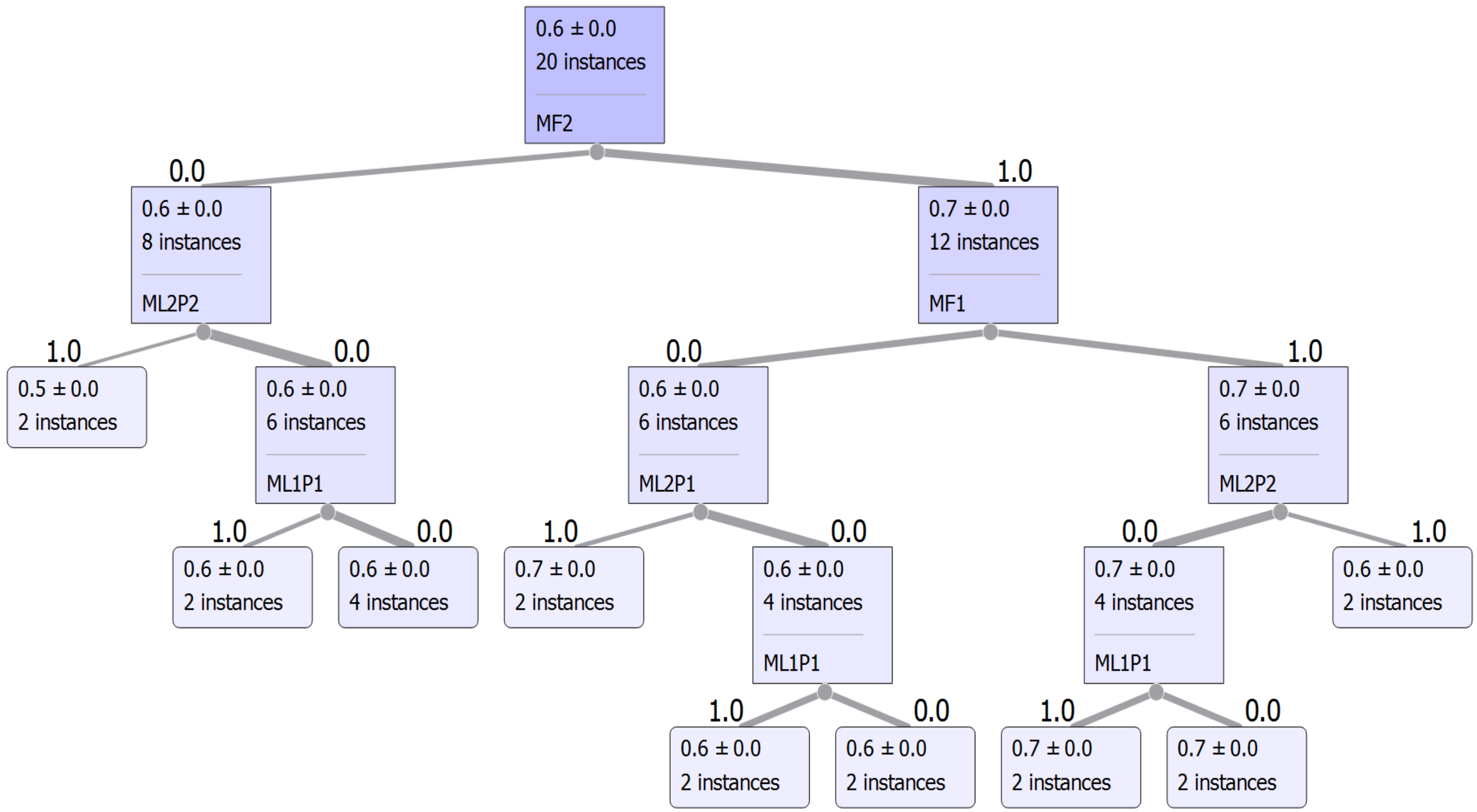


AI Engine
Visualization Engine
Graph Engine
Controller Engine
HCI Engine
TXT
TXT
JPG
JPG
TXT
TXT
ML Engine
TXT
GP
NB
DL
RF
DT
LR
Human Engine

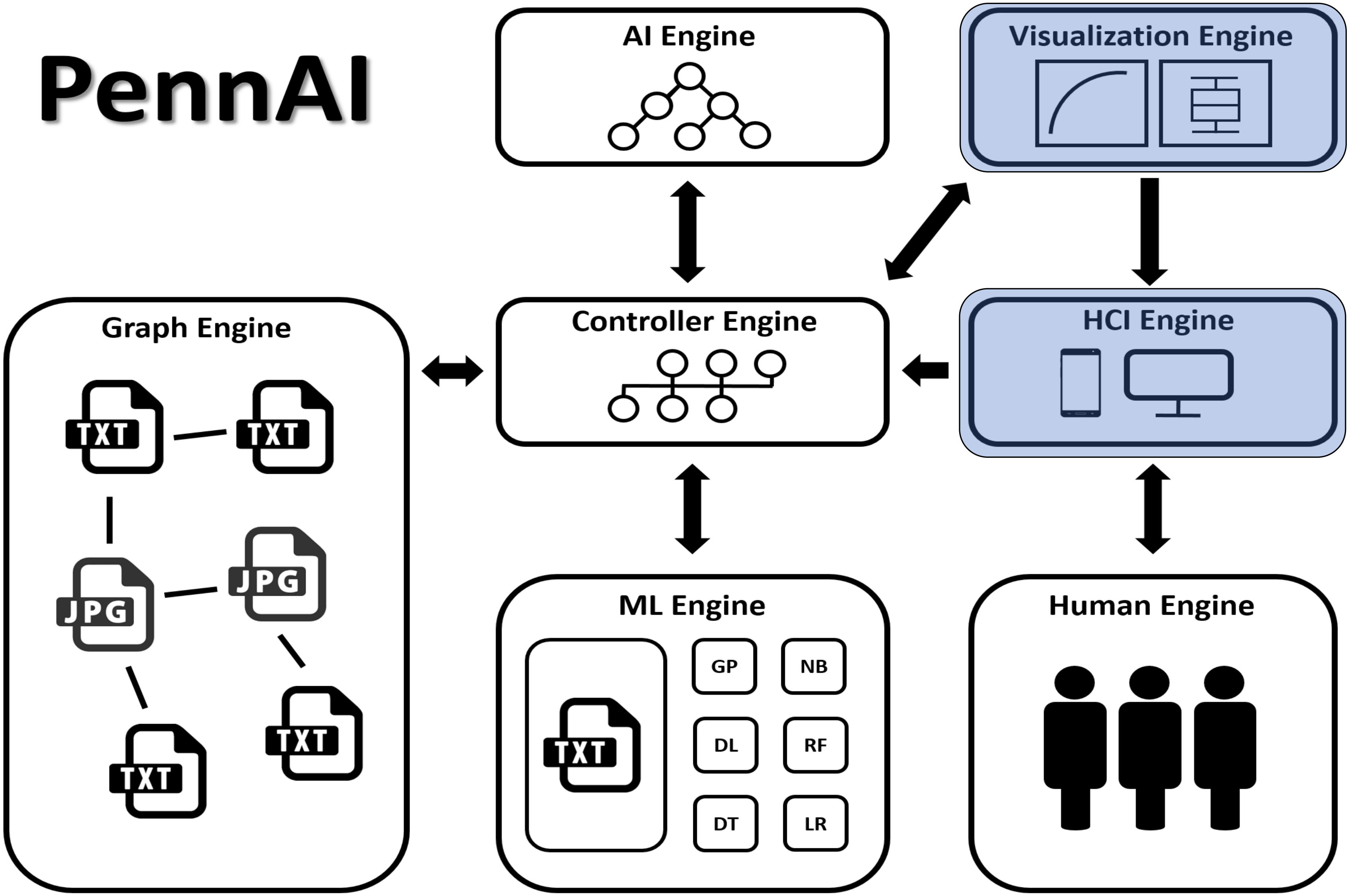
Machine learning results: Tabular form

Result	Dataset	MF1	MF2	ML1P1	ML1P2	ML2P1	ML2P2	AUC
1	1	1	0	1	0	0	0	0.56
2	1	1	0	0	1	0	0	0.52
3	1	1	0	0	0	1	0	0.61
4	1	1	0	0	0	0	1	0.55
5	2	1	1	0	0	0	1	0.69
6	2	1	1	0	1	0	0	0.72
7	2	1	1	1	0	0	0	0.69
8	3	0	1	1	0	0	0	0.57
9	3	0	1	0	1	0	0	0.67
10	3	0	1	0	0	1	0	0.68
11	4	1	0	0	0	1	0	0.56
12	4	1	0	0	1	0	0	0.65
13	4	1	0	1	0	0	0	0.67
14	4	1	0	0	0	0	1	0.53
15	5	1	1	0	0	0	1	0.53
16	5	1	1	0	1	0	0	0.66
17	5	1	1	1	0	0	0	0.78
18	6	0	1	1	0	0	0	0.59
19	6	0	1	0	1	0	0	0.58
20	6	0	1	0	0	1	0	0.67

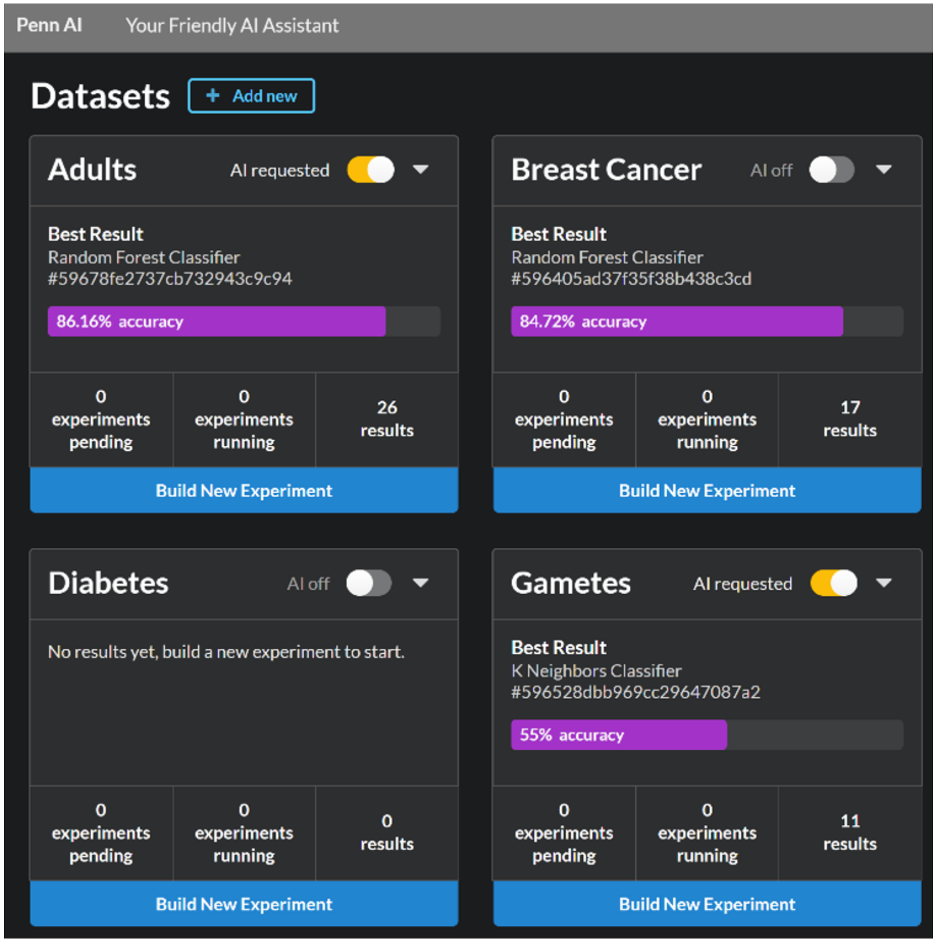
Machine learning results: Decision tree



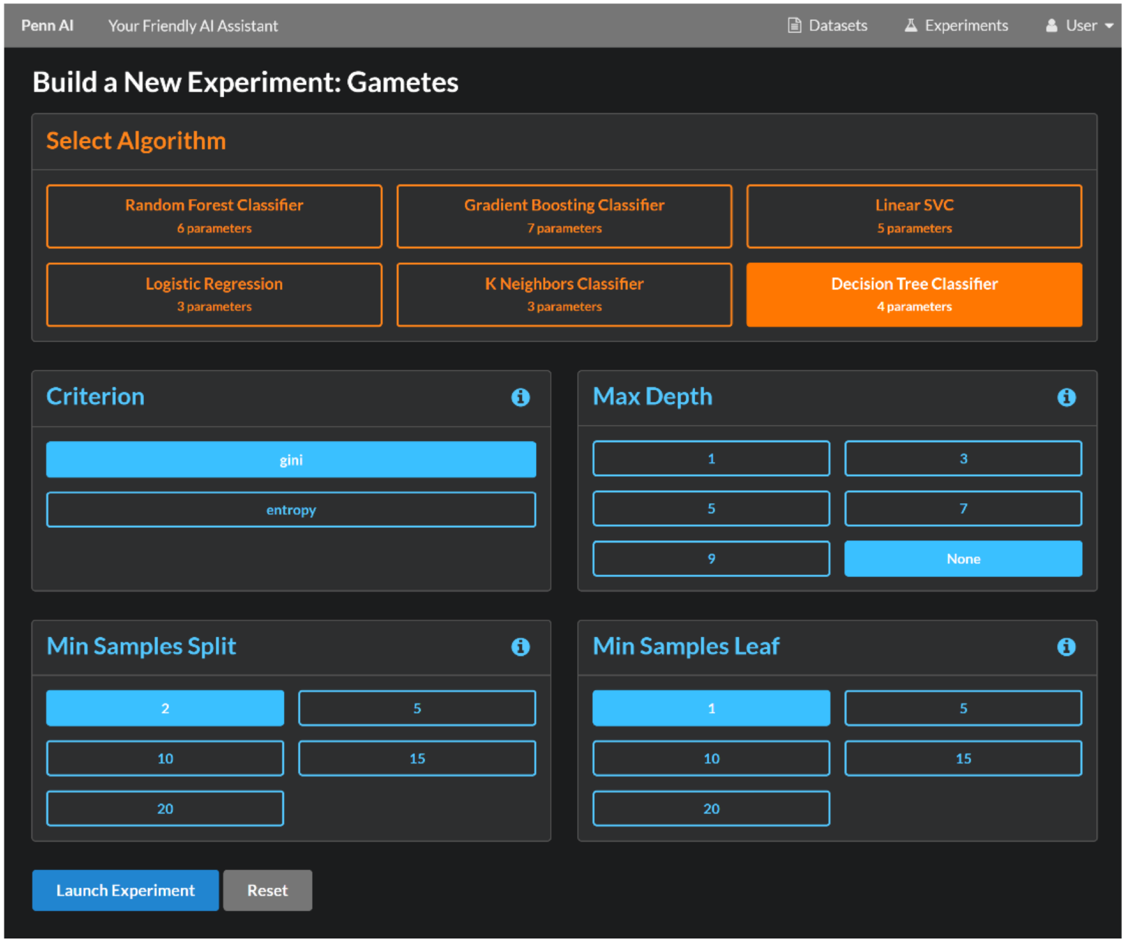
PennAI


AI Engine
Visualization Engine
Graph Engine
Controller Engine
HCI Engine
TXT
TXT
JPG
JPG
TXT
TXT
ML Engine
Human Engine
TXT
GP
NB
DL
RF
DT
LR

HCI: Dataset selection and history



HCI: Building an experiment



HCI/Visualization: Tabular results

Penn AI Your Friendly AI Assistant

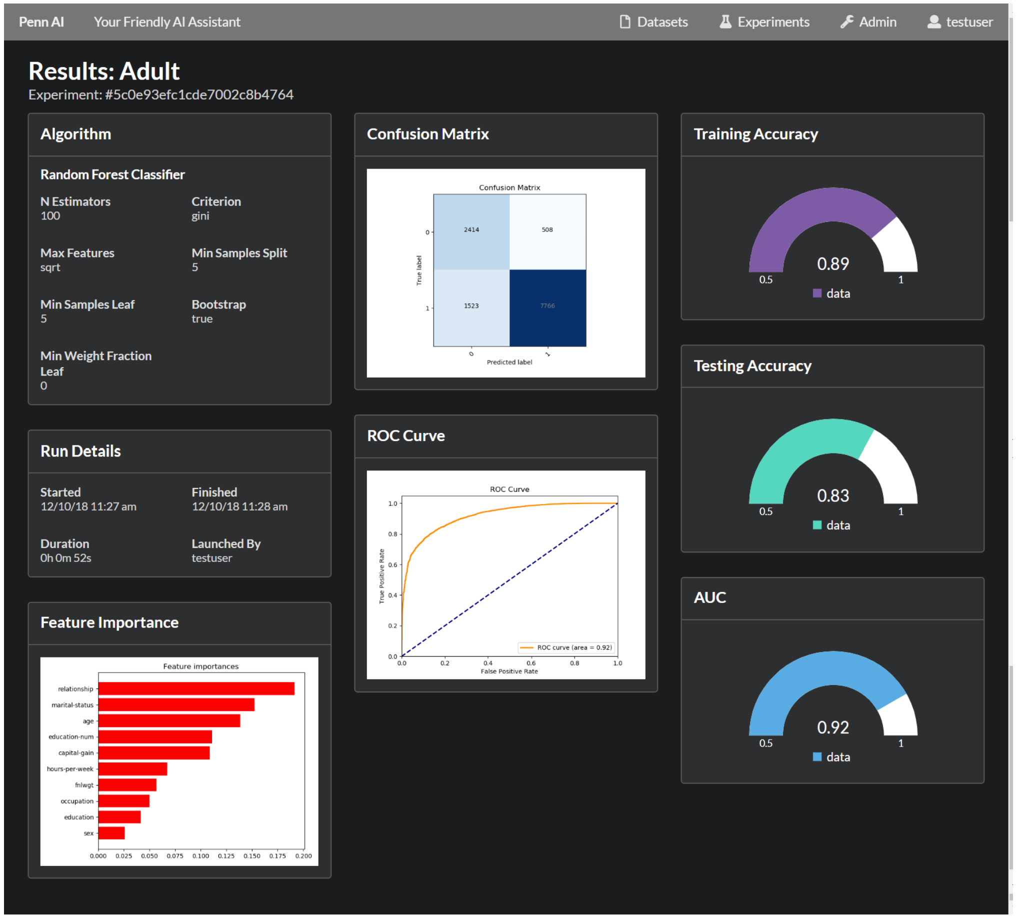
Datasets Experiments Admin testuser

Experiments

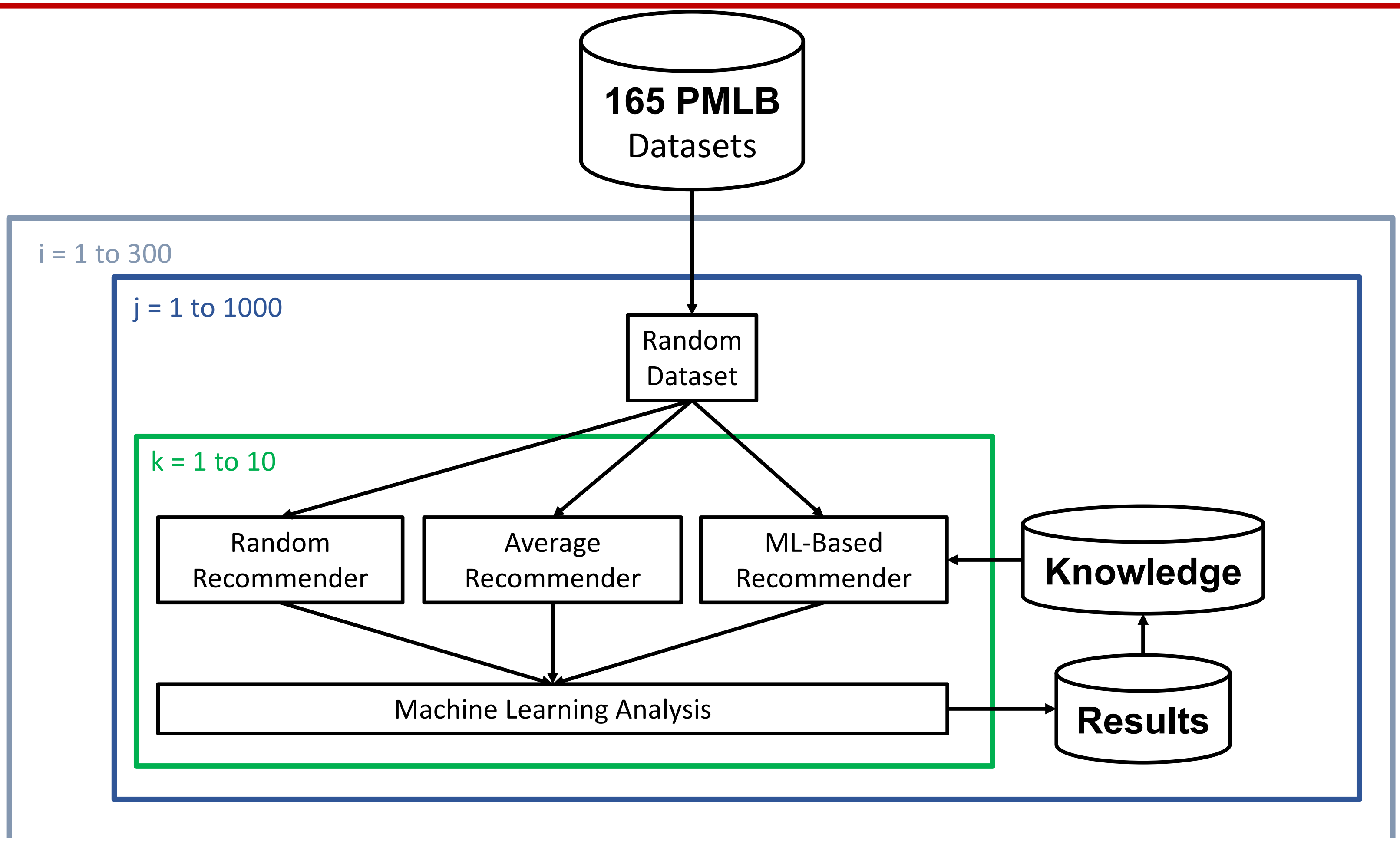
Status: Success Dataset: Adult Algorithm: All Reset filters **10 experiments**

Start Time	Accuracy	Dataset	Algorithm	Actions
12/10/18 11:36 am	0.83	Adult	Random Forest Classifier	
12/10/18 11:37 am	0.83	Adult	Random Forest Classifier	
12/10/18 12:27 pm	0.83	Adult	Random Forest Classifier	
12/10/18 11:27 am	0.83	Adult	Random Forest Classifier	
12/10/18 11:37 am	0.80	Adult	Gradient Boosting Classifier	
12/10/18 12:20 pm	0.79	Adult	Gradient Boosting Classifier	
12/10/18 12:29 pm	0.75	Adult	Gradient Boosting Classifier	
12/10/18 11:38 am	0.73	Adult	Gradient Boosting Classifier	
12/10/18 12:22 pm	0.72	Adult	Gradient Boosting Classifier	
12/10/18 12:20 pm	0.69	Adult	Gradient Boosting Classifier	

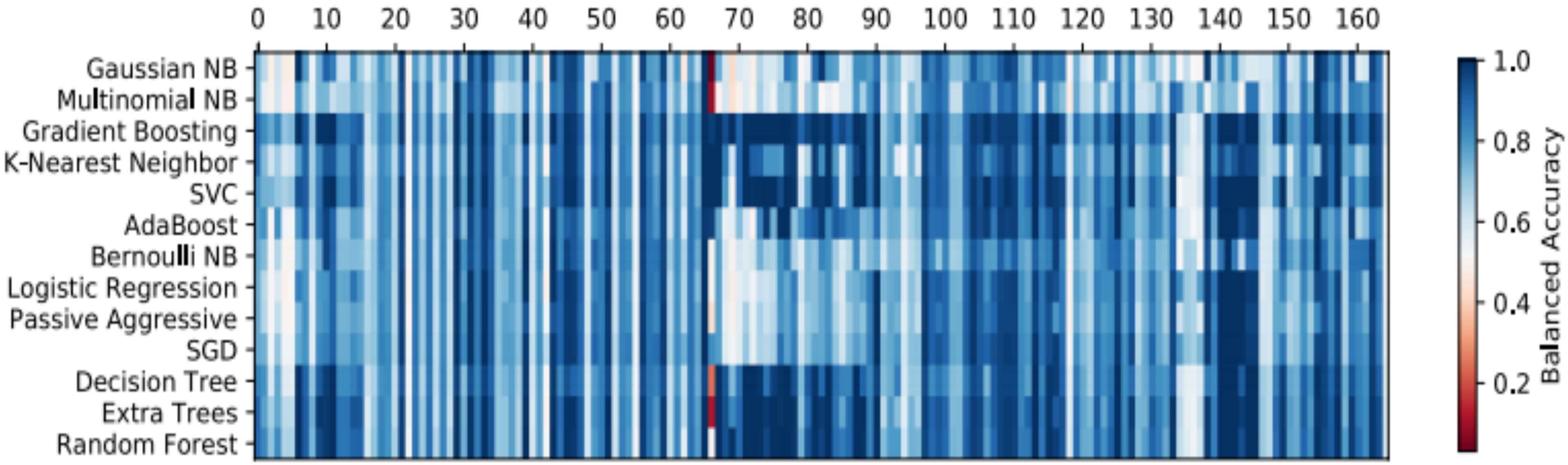
HCI/Visualization: Graphical results



PennAI evaluation: Schema



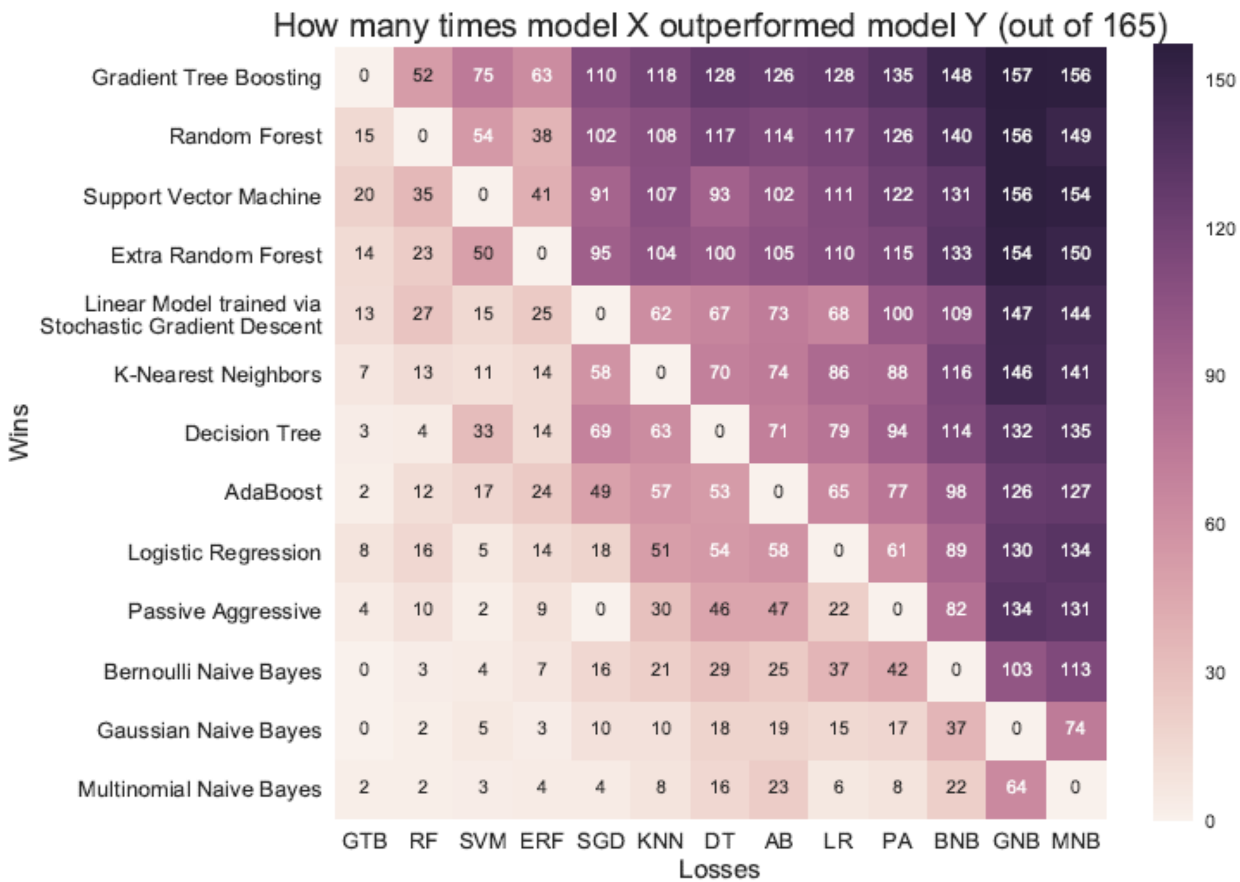
Results: 13 methods, 165 datasets



Penn Machine Learning Benchmarks (PMLB)

*BioData Min*ing 10:36 (2017)

Results: Tournament of the 13 methods



pennAI.org

pennai.org

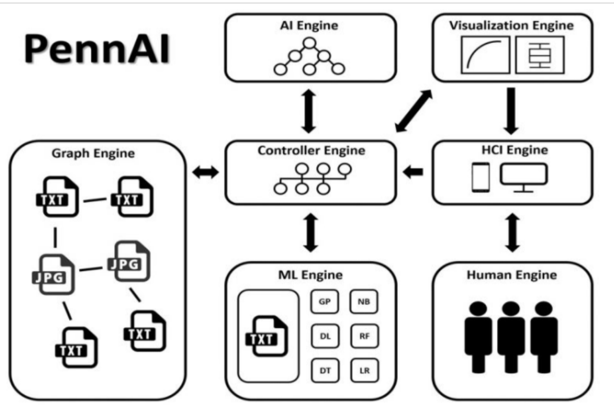
Penn AI

Accessible artificial intelligence from the University of Pennsylvania

HOME | GOAL | DEVELOPMENT TEAM | PUBLICATIONS | SOFTWARE | PRESS

Accessible Artificial Intelligence from Penn

We are pleased to launch this website dedicated to news and updates about PennAI – an accessible **artificial intelligence** system developed by at the University of Pennsylvania's **Perelman School of Medicine** by faculty, staff, and students from the **Penn Institute for Biomedical Informatics** (IBI). The components of PennAI include a human engine (i.e., the user); a user-friendly interface for interacting with the AI; a **machine learning** engine for data mining; a controller engine for launching jobs and keeping track of analytical results; a **graph database** for storing data and results (i.e., the memory); an AI engine for monitoring results and automatically launching or recommending new analyses; and a visualization engine to displaying results and analytical knowledge. This AI system provides a comprehensive set of integrated components for automated machine learning (**AutoML**), thus providing a data science assistant for generating useful results from large and complex data problems. More details can be found in our PennAI **publications**.



Accessible artificial intelligence

And as of this afternoon,
PennAI is now live!

https://github.com/EpistasisLab/pennai

Acknowledgments

- PennAI Team
  - Josh Cohen, Weixuan Fu, Paul Kopec, Bill La Cava, Randy Olson, Moshe Sipper, Sharon Tartarone, Heather Williams
- NIH grants R01s AI11679, LM012601, LM010098, UC4 DK112217

- jhmoore@upenn.edu
- epistasis.org, epistasisblog.org
- twitter.com: @moorejh

- PennAI.org

In summary…

- AI has a long tradition of explainability
  - But with newer machine learning methods, explainability has eluded us
- In order for AI to be usable by (not-always expert) researchers, the method and the results have to be explainable
  - Usability: Does "it" work?
- In order for AI to be useful to (not-always expert) researchers, the method and the results have to be explainable
  - Usefulness: Will it be used?

Thank you!

jhholmes@pennmedicine.upenn.edu