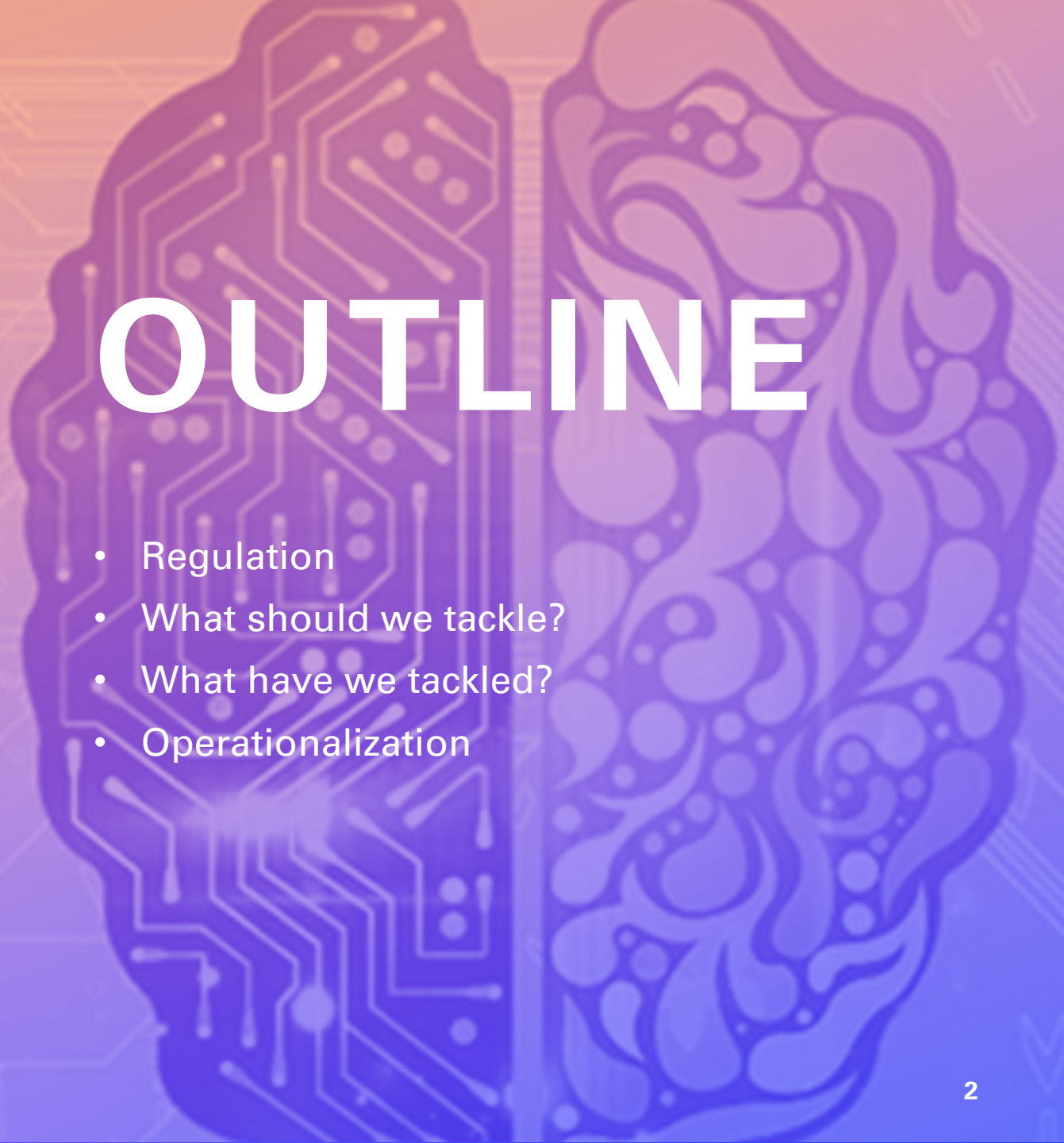
AI APPLICATIONS TO CLINICAL CHEMISTRY

Med Lab Week 2025, PLMS
Dan Holmes, MD FRCPC
Head, Department of Pathology and Laboratory Medicine
St. Paul's Hospital, Vancouver BC

MACHINE LEARNING



OUTLINE

- Regulation
- What should we tackle?
- What have we tackled?
- Operationalization

AI in Routine Clinical Care

- Background
  - AI applications in diagnostics are not new.
  - Hematology Analyzers from Sysmex use CellaVision's Neural Network Image analysis software for WBC classification.
  - This technology has been in clinical use since ~2001.
- So what is new?
  - There are many varieties of AI/ML now
  - There are many relatively easy-to-use (even low code/ no code) tools for building ML.

FDA AI/ML List

Artificial Intelligence and Machine Learning (AI/ML)-Enabled Medical Devices

Share | Post | Linkedin | Email | Print

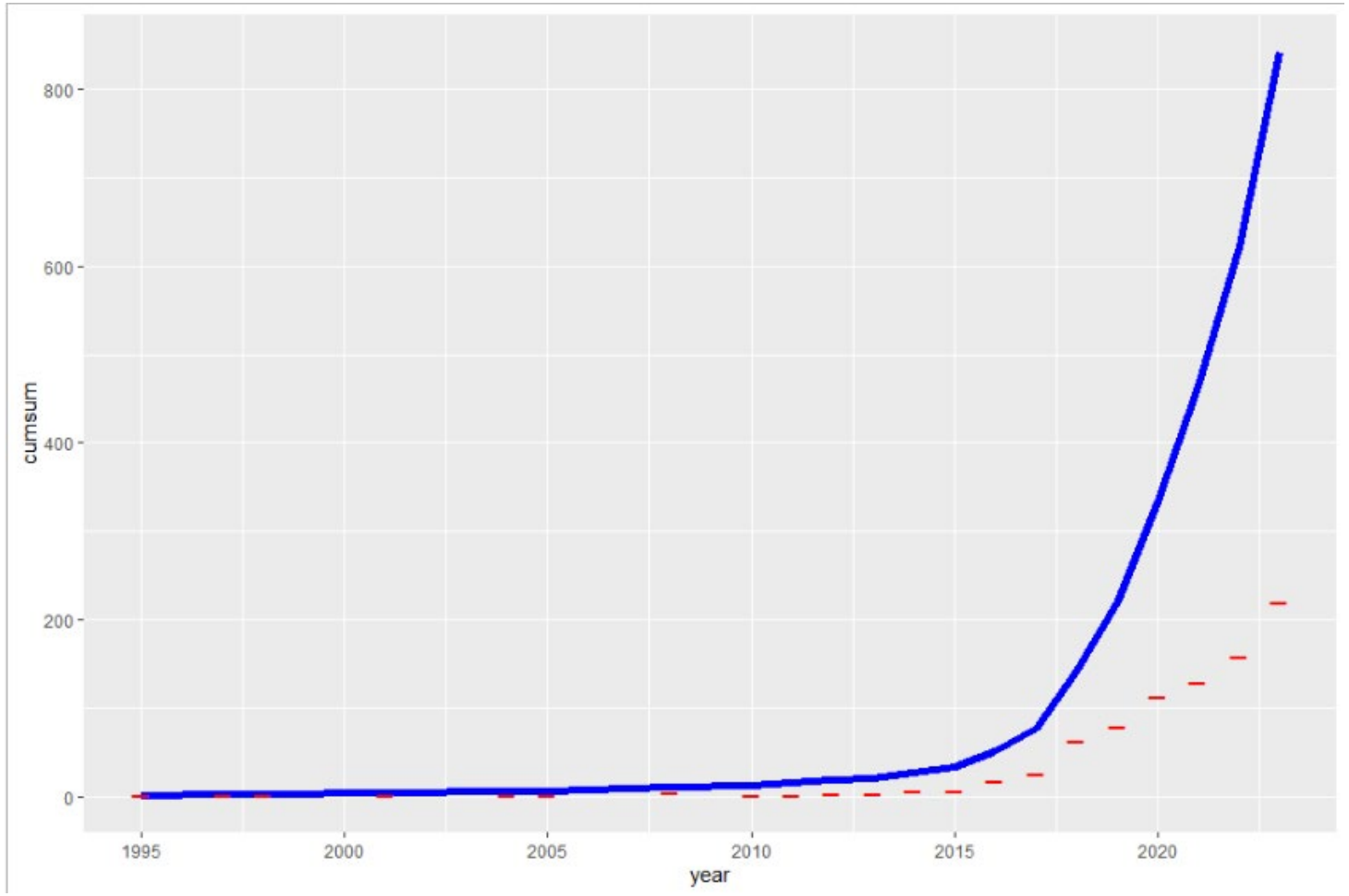
August 7, 2024 update: The U.S. Food and Administration updated the list of Artificial Intelligence and Machine Learning (AI/ML)-Enabled Medical Devices. With this update, the FDA has authorized 950 AI/ML-enabled medical devices.

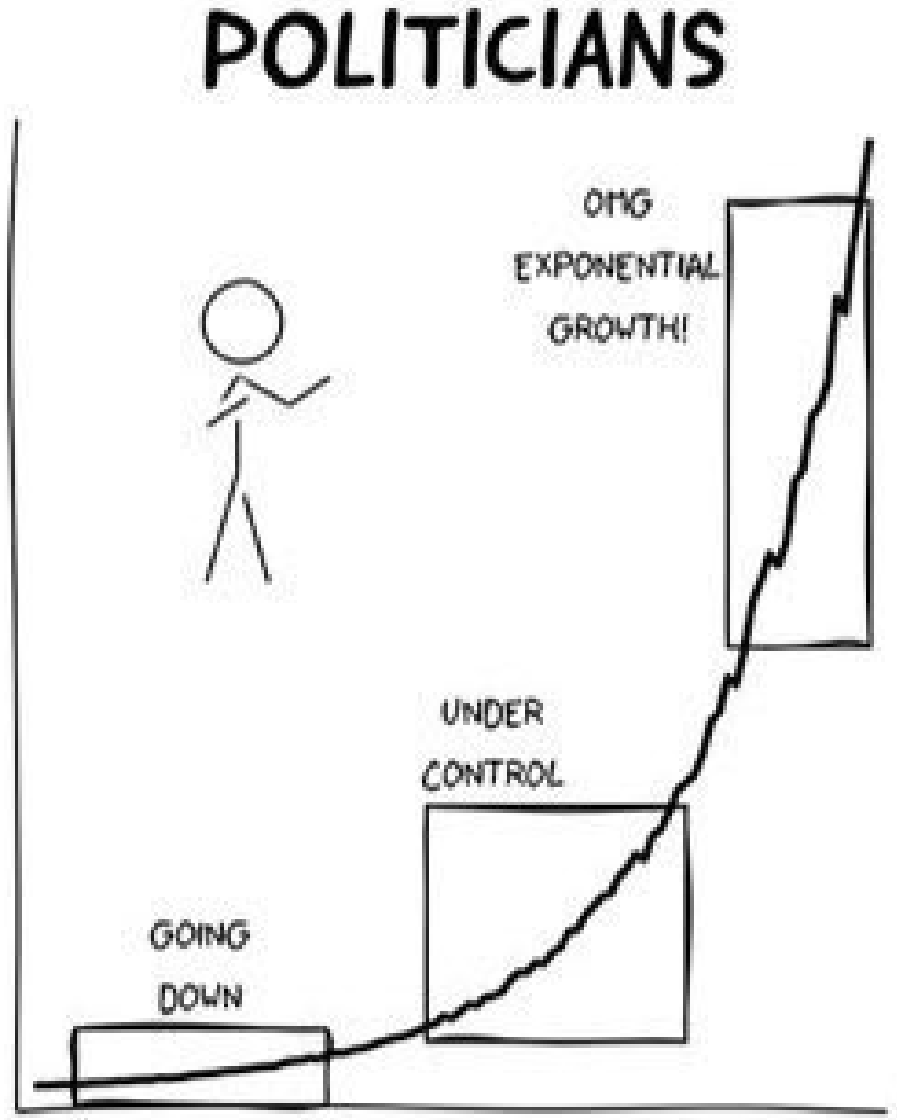
The FDA is providing this list of AI/ML-enabled medical devices marketed in the United States as a resource to the public about these devices and the FDA's work in this area. The devices in this list have met the FDA's applicable premarket requirements, including a focused review of the devices' overall safety and effectiveness, which includes an evaluation of appropriate study diversity based on the device's intended use and technological characteristics.

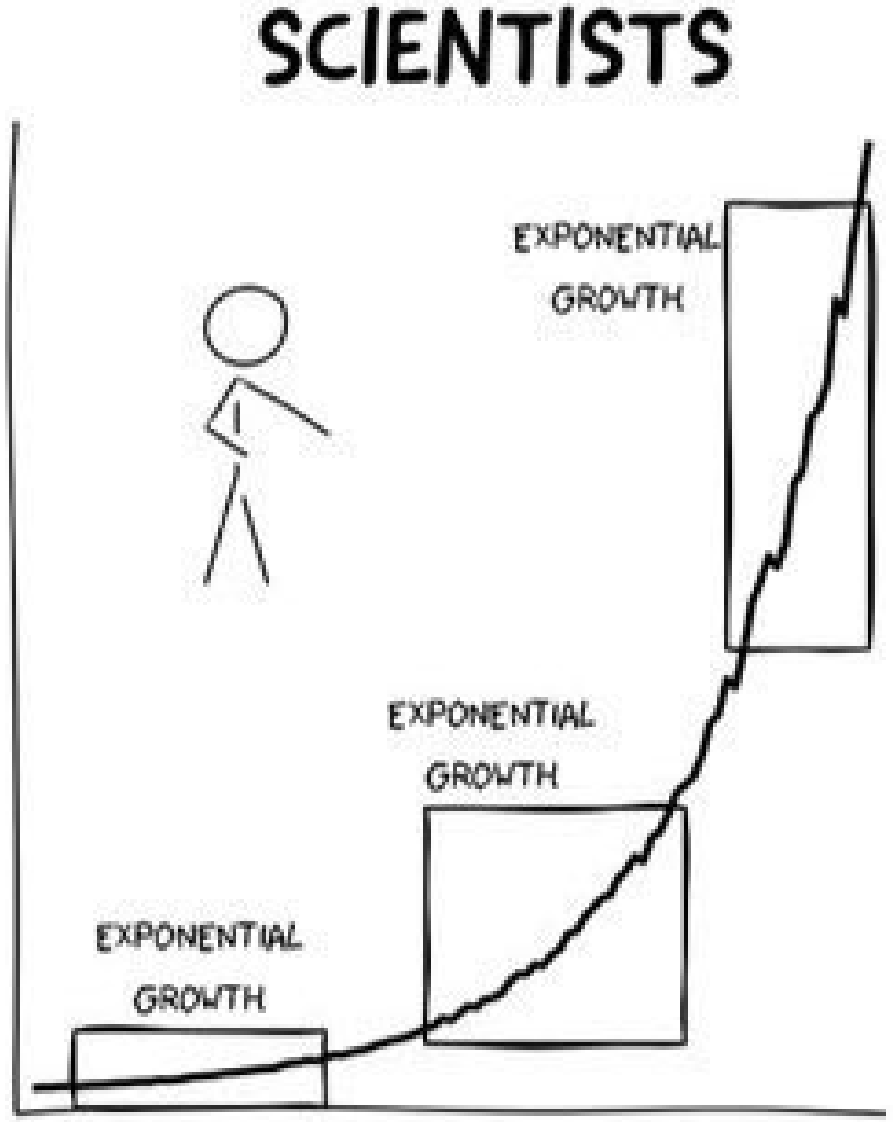
- URL: https://www.fda.gov/medical-devices/software-medical-device-samd/artificial-intelligence-and-machine-learning-aiml-enabled-medical-devices
- Lists 1017 devices on 14-Apr-2025

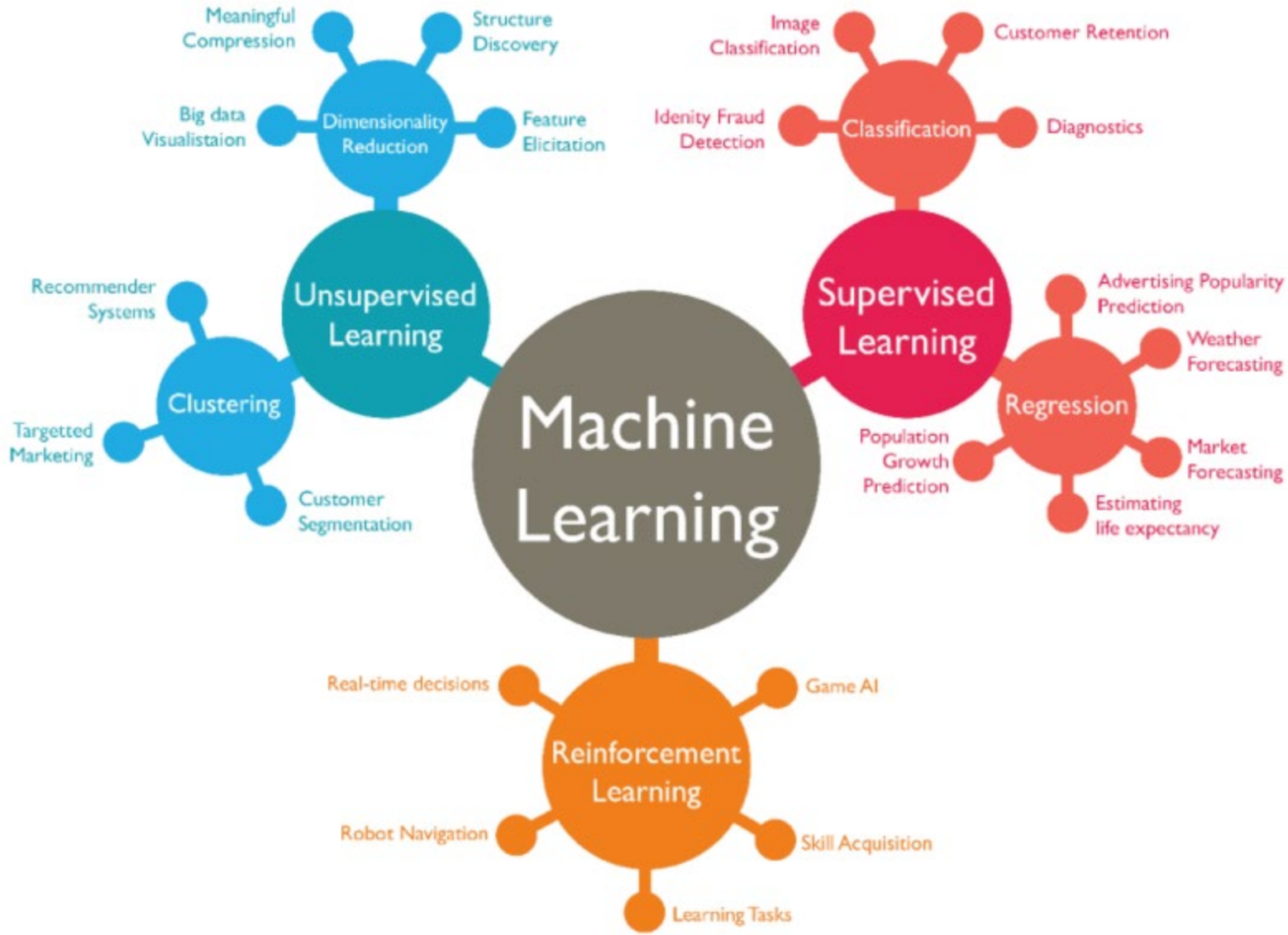
FDA ML/AI registrations

- The availability of new methods easy-to-use tools has driven the FDA registrations
- Blue line is cumulative FDA registrations
- Red dashes are annual registrations




POLITICIANS
OMG
EXPONENTIAL
GROWTH!
UNDER
CONTROL
GOING
DOWN


SCIENTISTS
EXPONENTIAL
GROWTH
EXPONENTIAL
GROWTH
EXPONENTIAL
GROWTH



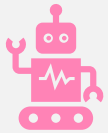
https://www.linkedin.com/pulse/artificial-intelligence-5-taxonomy-machine-learning-deep-ajit-jaokar/

Why the Proliferation?

The open-source tools for Data Science and Analytics have become very mature in the past 5 years with a massive user base:

R: R-Studio Tidyverse tools

Python: Pandas, SciPi, Numpy, Matplotlib



Low barrier of entry Machine Learning, AI and Deep-learning toolkits have been developed

R: Caret package, Tidymodels

Python: Keras, Tensorflow, Pytorch, Scikit-learn

The Ideal Task


Tell me the truth...I'm...I'm ready to hear it.
YOU DON'T NEED MACHINE LEARNING FOR THAT
imgflip.com

Rachael Tatman
@rctatman

Replying to @almeezyuh

1. You have a very repetitive task that you have already done by hand a lot and want to automate
2. The task is fairly easy to teach a human to do (let's say < 5 minutes)
3. As long as you mostly do the task right you'll be fine (NOT high stakes)

11:59 AM · Jan 31, 2022 · Twitter Web App

Diagnosis vs Decision Support

- We often focus on the more glamorous, high-stakes problem of diagnostic classification.
- We can also address lower-stakes problems of decision support in repetitive, time consuming non-diagnostic tasks.
- e.g. peak quality evaluation, interference detection, filtering false-positive peaks.

Health Canada AI and SaMD

- https://www.canada.ca/en/health-canada/services/drugs-health-products/medical-devices/application-information/guidance-documents/pre-market-guidance-machine-learning-enabled-medical-devices.html

- https://www.canada.ca/en/health-canada/services/drugs-health-products/medical-devices/application-information/guidance-documents/software-medical-device-guidance-document.html

Examples

Supervised Learning Problems

- Mostly *classification* problems in Clinical Chemistry
- I have multiple known analytes (“features”) and I have diagnoses (“labels”) but I want to improve on my rules-based algorithm.
  - Prenatal Screening
  - Newborn Screening
  - Steroid metabolomes
- I have more features than I can easily evaluate and no a priori method to evaluate, but I have labels by another means.
  - High plex or untargeted proteomics/metabolomics/lipidomics data on cases with known diagnosis (e.g. by histology).

Urine Steroid Metabolomics as a Novel Tool for Detection of Recurrent Adrenocortical Carcinoma

Vasileios Chortis,[1,2,3] Irina Bancos,[1,4] Thomas Nijman,[5] Lorna C. Gilligan,[1] Angela E. Taylor,[1] Cristina L. Ronchi,[1,2,3,6] Michael W. O'Reilly,[1,2,3] Jochen Schreiner,[6] Miriam Asia,[2,3] Anna Riester,[7] Paola Perotti,[8] Rosella Libé,[9] Marcus Quinkler,[10] Letizia Canu,[11] Isabel Paiva,[12] Maria J. Bugalho,[13] Darko Kastelan,[14] M. Conall Dennedy,[15] Mark Sherlock,[16] Urszula Ambroziak,[17] Dimitra Vassiliadi,[18] Jerome Bertherat,[9] Felix Beuschlein,[7,19] Martin Fassnacht,[6,20,21] Jonathan J. Deeks,[22,23] Michael Bi[illegible][5] and Wiebke Ar[illegible][1,2,3,23]

ADRENAL CA RECURRENCE

[1]Institute of Metabolism and Systems Research, University of Birmingham, Birmingham, UK;
[2]Centre for Endocrinology, Diabetes and Metabolism, Birmingham Health Partners, Birmingham

Clinical Chemistry 64:11
000–000 (2018)

Using Machine Learning to Aid the Interpretation of Urine Steroid Profiles

Edmund H. Wilkes,[1] Gill Rumsby,[1] and Gary M. Woodward[1*]

BACKGROUND: Urine steroid profiles are used in clinical practice for the diagnosis and monitoring of disorders of steroidogenesis and adrenal pathologies. Machine learning (ML) algorithms are powerful computational tools used extensively for the recognition of patterns in large data sets. Here, we investigated the utility of various ML algorithms for the automated biochemical interpretation of urine steroid profiles to support current clinical practices.

knowledge. Testing procedures that produce multiple analyte measurements within a single test represent an additional level of complexity. An example of this is urine steroid profiling (USP)[2], a test used for the diagnosis and monitoring of disorders of steroidogenesis and adrenal pathologies. The interpretation of USPs requires the integration of multiple discrete steroid measurements in the context of specific diagnostic questions. This requires a high degree of specialist clinical and technical knowl-

URINE STEROID DECISION SUPPORT

JAMA Network | Open™

Original Investigation | Pathology and Laboratory Medicine

Use of Steroid Profiling Combined With Machine Learning for Identification and Subtype Classification in Primary Aldosteronism

Graeme Eisenhofer, PhD; Claudio Durán, MS; Carlo Vittorio Cannistraci, PhD; Mirko Peitzsch, PhD; Tracy Ann Williams, PhD; Anna Riester, MD; Jacopo Burrello, MD; Fabrizio Buffolo, MD; Aleksander Prejbisz, MD; Felix Beuschlein, MD; Andrzej Januszewicz, MD; Paolo Mulatero, MD; Jacques W. M. Lenders, MD; Martin Reincke, MD

Abstract

IMPORTANCE Most patients with primary aldosteronism, a major cause of secondary hypertension, are not identified or appropriately treated because of difficulties in diagnosis and subtype classification. Applications of artificial intelligence combined with mass spectrometry–based steroid profiling could address this problem.

OBJECTIVE To assess whether plasma steroid profiling combined with machine learning might facilitate diagnosis and treatment stratification of primary aldosteronism patients with unilateral adenomas due to pathogenic *KCNJ5* sequence variants.

Key Points

Question Does steroid profiling combined with machine learning offer a potential 1-step strategy to facilitate diagnosis and subtype classification for treatment stratification of patients with primary aldosteronism?

Findings This diagnostic study involving patients tested for primary aldosteronism found that those with

PRIMARY ALDOSTERONISM DX, SUBTYPE CLASSIFICATION

Clinical Chemistry 66:9
1210–1218 (2020)

A Machine Learning Approach for the Automated Interpretation of Plasma Amino Acid Profiles

Edmund H. Wilkes,[a] Erin Emmett,[b] Luisa Beltran,[b] Gary M. Woodward,[c] and Rachel S. Carling[b,d]*

BACKGROUND: Plasma amino acid (PAA) profiles are used in routine clinical practice for the diagnosis and monitoring of inherited disorders of amino acid metabolism, organic acidemias, and urea cycle defects. Interpretation of PAA profiles is complex and requires substantial training and expertise to perform. Given previous demonstrations of the ability of machine learning (ML) algorithms to interpret complex clinical biochemistry data, we sought to

profiles. This would be particularly useful in labora with limited resources and large workloads. We pr the necessary code for other laboratories to develop own decision support tools.

Introduction

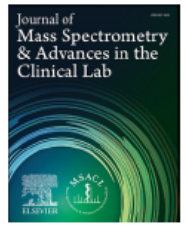
NBS DECISION SUPPORT

Same Data – Tutorial

Contents lists available at ScienceDirect

Journal of Mass Spectrometry and Advances in the Clinical Lab

journal homepage: www.sciencedirect.com/journal/journal-of-mass-spectrometry-and-advances-in-the-clinical-lab



Check for updates

Supervised machine learning in the mass spectrometry laboratory: A tutorial

Edward S. Lee [a,b], Thomas J.S. Durant [a,b,*]

[a] Department of Laboratory Medicine, at Yale School of Medicine, New Haven, CT, USA
[b] Department of Laboratory Medicine, at Yale New Haven Hospital, New Haven, CT, USA

ARTICLE INFO

Keywords:
Supervised machine learning
Xgboost
Artificial intelligence
Mass spectrometry
Amino acid

ABSTRACT

As the demand for laboratory testing by mass spectrometry increases, so does the need for automated methods for data analysis. Clinical mass spectrometry (MS) data is particularly well-suited for machine learning (ML) methods, which deal nicely with structured and discrete data elements. The alignment of these two fields offers a promising synergy that can be used to optimize workflows, improve result quality, and enhance our understanding of high-dimensional datasets and their inherent relationship with disease. In recent years, there has been an increasing number of publications that examine the capabilities of ML-based software in the context of chromatography and MS. However, given the historically distant nature between the fields of clinical chemistry and computer science, there is an opportunity to improve technological literacy of ML-based software within the clinical laboratory scientist community. To this end, we present a basic overview of ML and a tutorial of an ML-based experiment using a previously published MS dataset. The purpose of this paper is to describe the fundamental principles of supervised ML, outline the steps that are classically involved in an ML-based experiment, and discuss the purpose of good ML practice in the context of a binary MS classification problem.

Clinica Chimica Acta 567 (2025) 120086

Contents lists available at ScienceDirect

Clinica Chimica Acta

journal homepage: www.elsevier.com/locate/cca

Research Paper

Artificial intelligence aided serum protein electrophoresis analysis o Finnish patient samples: Retrospective validation

Tapio Lahtiharju [a,*], Lassi Paavolainen [b], Janne Suvisaari [a], Pasi Nokelainen [a], Emmi R Mikko Anttonen [a], Outi Itkonen [a]

ELECTROPHORESIS INTERPRETATION

Streamlining Quality Review of Mass Spectrometry Data in the Clinical Laboratory by Use of Machine Learning

Min Yu, MD, PhD; Lindsay A. L. Bazydlo, PhD; David E. Bruns, MD; James H. Harrison Jr, MD, PhD

• *Context.*—Turnaround time and productivity of clinical mass spectrometric (MS) testing are hampered by time-consuming manual review of the analytical quality of MS data before release of patient results.

Objective.—To determine whether a classification model created by using standard machine learning algorithms can verify analytically acceptable MS results and thereby reduce manual review requirements.

Design.—We obtained retrospective data from gas chromatography–MS analyses of 11-nor-9-carboxy-delta-9-tetrahydrocannabinol (THC-COOH) in 1267 urine samples. The data for each sample had been labeled previously as either analytically unacceptable or acceptable by manual review. The dataset was randomly split into training and test sets (848 and 419 samples, respectively), maintaining equal proportions of acceptable (90%) and unacceptable (10%) results in each set. We stratified 10-fold cross-validation in assessing the abilities of 6 supervised machine learning algorithms to distinguish unacceptable from acceptable assay results in the training dataset. The classifier with the highest recall was used to build a final model, and its performance was evaluated against the test dataset.

Results.—In comparison testing of the 6 classifiers, a model based on the Support Vector Machines algorithm yielded the highest recall and acceptable precision. After optimization, this model correctly identified all unacceptable results in the test dataset (100% recall) with a precision of 81%.

Conclusions.—Automated data review identified all analytically unacceptable assays in the test dataset, while reducing the manual review requirement by about 87%. This automation strategy can focus manual review only on assays likely to be problematic, allowing improved throughput and efficiency without reducing quality.

(*Arch Pathol Lab Med.* 2019;143:990–998; doi: 10.5858/arpa.2018-0238-OA)

PEAK QUALITY REVIEW DECISION SUPPORT

Clin Chem Lab Med 2021; 59(8): 1392–1399 DE GRUYTER

Gabriel L. Streun, Andrea E. Steuer, Lars C. Ebert, Akos Dobay and Thomas Kraemer*

Interpretable machine learning model to detect chemically adulterated urine samples analyzed by high resolution mass spectrometry

https://doi.org/10.1515/cclm-2021-0010
Received January 4, 2021; accepted March 5, 2021;
published online March 22, 2021

Abstract

Objectives: Urine sample manipulation including substitution, dilution, and chemical adulteration is a continuing challenge for workplace drug testing, abstinence control, and doping control laboratories. The simultaneous detection of sample m[illegible] one single analytical measurement would be highly advantageous. Machine learning algorithms are able to learn from existing datasets and predict outcomes of new

Results: Following 10-fold cross-validation, the mean sensitivity, specificity, positive predictive value, and negative predictive value was 88.9, 92.0, 91.9, and 89.2%, respectively. A diverse test set (n=202) containing treated and untreated urine samples could be correctly classified with an accuracy of 95.4%. In addition, 14 important features and four potential biomarkers were extracted.

Conclusions: With interpretable retention time aligned liquid chromatography high-resolution mass spectrometry [illegible] could be established [illegible] manipulation. The incorporation of our model into routine clinical or forensic analysis allows simultaneous LC-MS analysis and

SAMPLE ADULTERATION

Clinical Chemistry 65:5
674–683 (2019)

Performance of the MasSpec Pen for Rapid Diagnosis of Ovarian Cancer

Marta Sans,[1†] Jialing Zhang,[1†] John Q. Lin,[1] Clara L. Feider,[1] Noah Giese,[1] Michael T. Breen,[2] Katherine Sebastian,[3] Jinsong Liu,[4] Anil K. Sood,[5] and Livia S. Eberlin[1*]

BACKGROUND: Accurate tissue diagnosis during ovarian cancer surgery is critical to maximize cancer excision and define treatment options. Yet, current methods for intraoperative tissue evaluation can be time intensive and subjective. We have developed a handheld and biocompatible device coupled to a mass spectrometer, the MasSpec Pen, which use[illegible] extraction and [illegible] the performance of this technology for ovarian cancer [illegible] serous cancer prediction and thus has potential for clinical use for rapid and accurate ovarian cancer diagnosis.

Ovarian cancer is a highly lethal disease and the fifth [illegible] in women [illegible]. Accurate diagnosis and stratification of ovarian cancer is important to develop personalized treatment

INTRAOPERATIVE MASS SPEC

St John *et al. Breast Cancer Research* (2017) 19:59
DOI 10.1186/s13058-017-0845-2

Breast Cancer Research

RESEARCH ARTICLE Open Access

Rapid evaporative ionisation mass spectrometry of electrosurgical vapours for the identification of breast pathology: towards an intelligent knife for breast cancer surgery

Edward R. St John[1], Julia Balog[2,3], James S. McKenzie[2], Merja Rossi[2], April Covington[1], Laura Muirhead[1], Zsolt Bodai[2], Francesca Rosini[2,4], Abigail V. M. Speller[2,4], Sami Shousha[4], Rathi Ramakrishnan[4], Ara Darzi[1], Zoltan Takats[2] and Daniel R. Leff[1,5]

INTRAOPERATIVE MASS SPEC

Delta check – Machine Learning

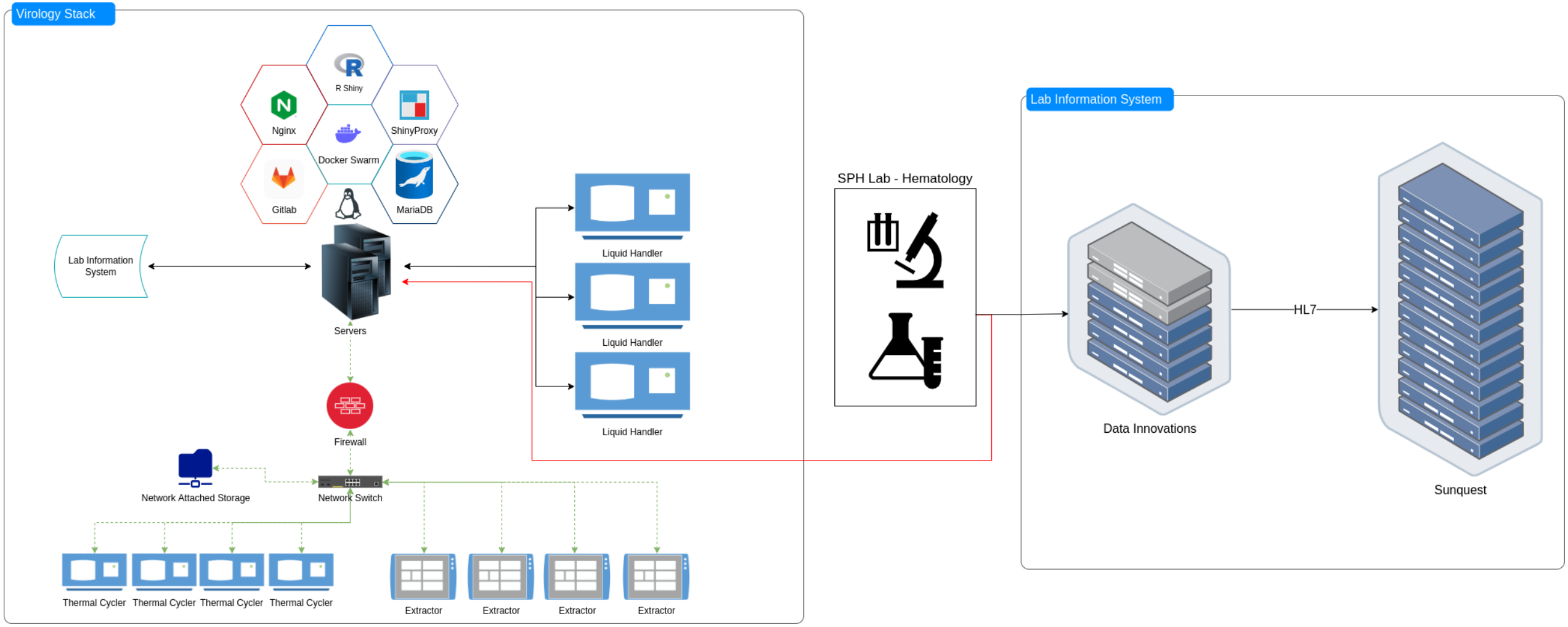
Data & Model

- Five years of historical data, SPH, MSJ
  - "Complete Blood Count(CBC)", "WBC Differential(DIF)", "CBC and Differential(CBCDIF)", "CBC & Differential(CBCD)“
  - Simulated mislabeling and dilution errors as training & testing data
  - 65 features selected based on numerous experiments
- Extensive model evaluations: RF, EBM, CatBoost, LightGBM … , XGBoost
- Containerized pipeline components for data preprocessing, training, inference, explainability, and postprocessing steps
- Parameterized RMarkdown notebooks for development and documentations

Deployment

- Rshiny, Inotifywait, Plumber, pySocket, Docker Swarm
- MLOps system using KubeFlow (provider agnostic)

Delta check – on-premises



inical Chemistry 70:2
4–452 (2024)

Automating the Detection of IV Fluid Contamination Using Unsupervised Machine Learning

Nicholas C. Spies,[a,*] Zita Hubler,[a] Vahid Azimi,[a] Ray Zhang,[b] Ronald Jackups, Jr.,[a] Ann M. Gronowski,[a] Christopher W. Farnsworth,[a] and Mark A. Zaydman[a]

CKGROUND: Intravenous (IV) fluid contamination a common cause of [illegible] lay or misguide treatment decisions, leading to tient harm. Current approaches for detecting contam-

CONCLUSIONS: Accurate and automatable detection [illegible] BMP results is achiev without curating expertly labeled training data.

IV FLUID CONTAMINATION

Deployment


TO SURVIVE IN THE WILD, YOU NEED A RELIABLE SOURCE OF WATER
I.T. IS GONNA BLOCK THE DEPLOYMENT OF YOUR AI
SLURP

Operationalization Challenges

- Operationalization may be more work than development.
- Bullet-proofing and useability
- On premises deployment vs cloud deployment
- Infrastructure and sustainment
- Analytical hardware turnover/change
- Interoperability challenges

Conclusions

- Open-source ML and Data Science toolkits have become very mature and have lowered the barrier of entry substantially.
- Problems we are attracted to tend to be our higher stakes problems with some regulatory implications.
  - Deployment of models is very challenging.
  - Models may catastrophically break if input data changes in some way.
- Many of our vexing problems are simple and operational in nature and can be addressed with the same tools – often without ML.

Questions

contact:
dtholmes@mail.ubc.ca