Machine Learning—Examples from Particle Accelerators

Y. K. Wu

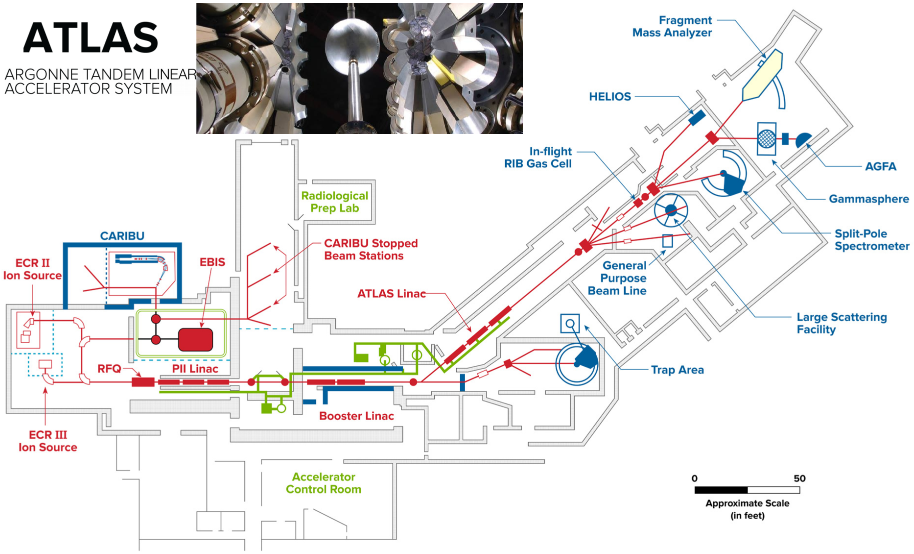
TUNL and Duke University

January 30, 2020

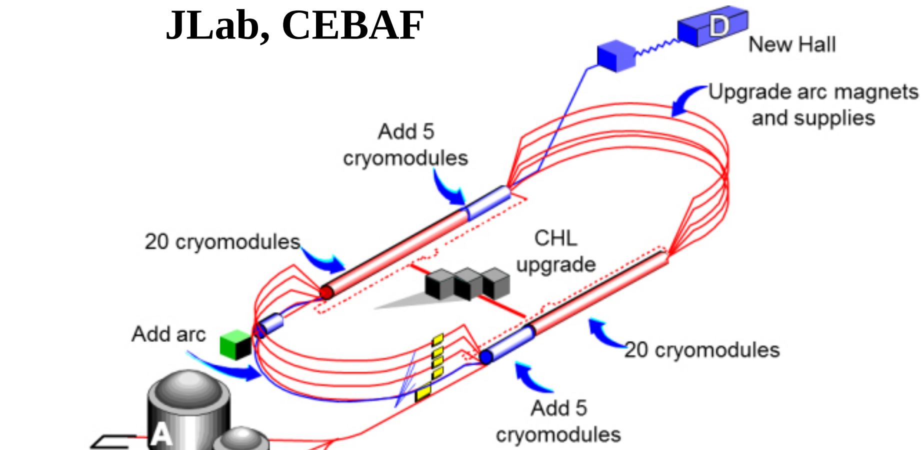
Work supported by **U.S. DOE Grant: DE-FG02-97ER41033**

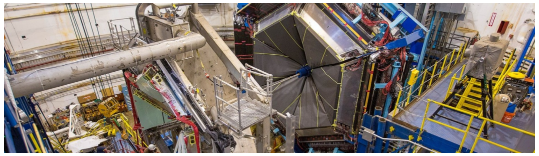
DoE Nuclear Physics User Facilities

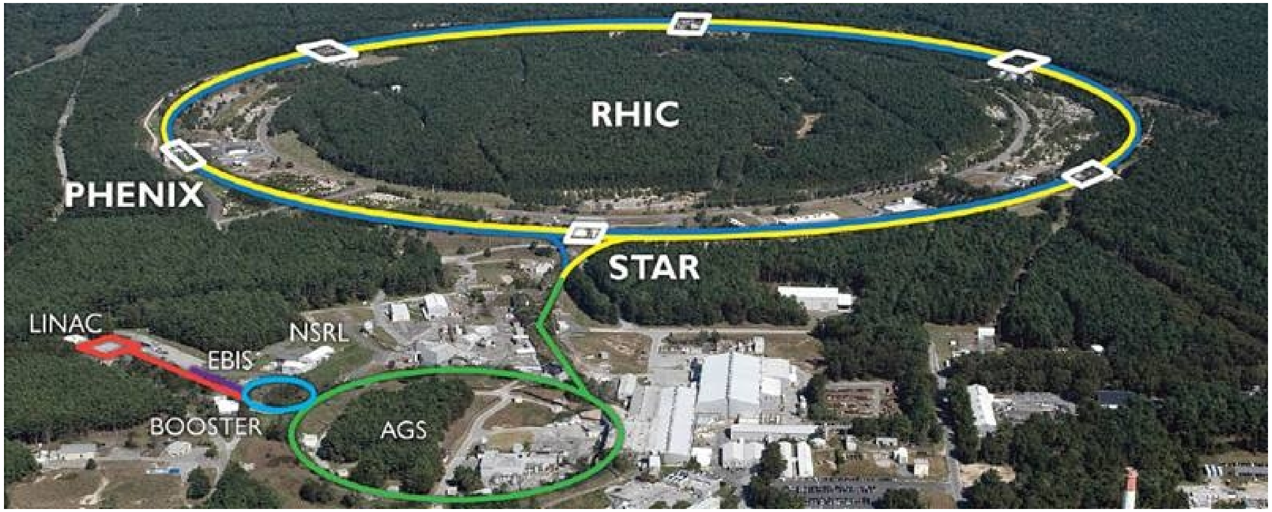
ATLAS
ARGONNE TANDEM LINEAR ACCELERATOR SYSTEM



JLab, CEBAF





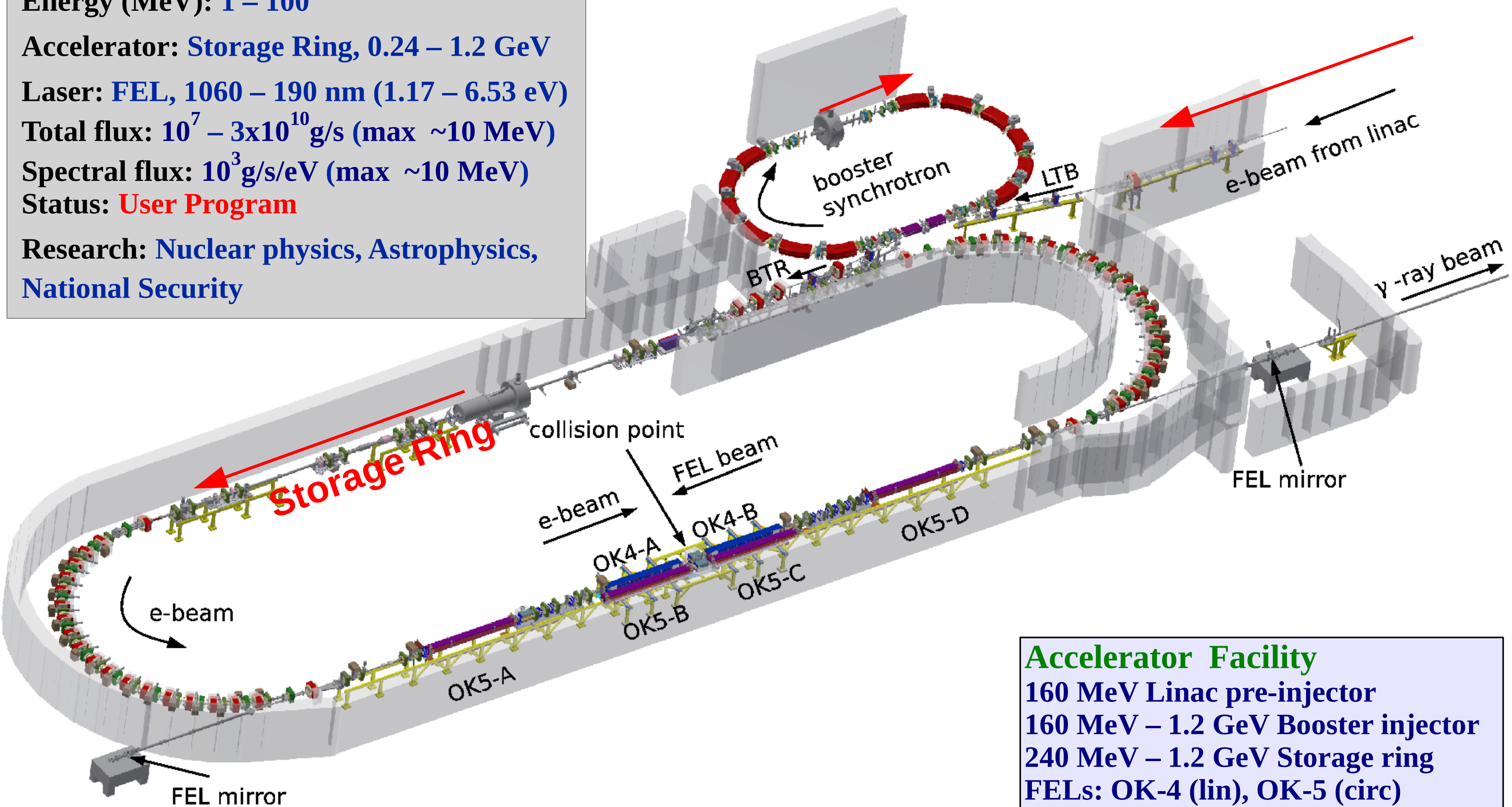


https://science.osti.gov/np/Facilities/User-Facilities
https://www.anl.gov/sites/www/files/2018-12/ATLAS_floor_plan_Dec2018.pdf
https://www.jlab.org/physics

HIGS/TUNL: Accelerator Facility

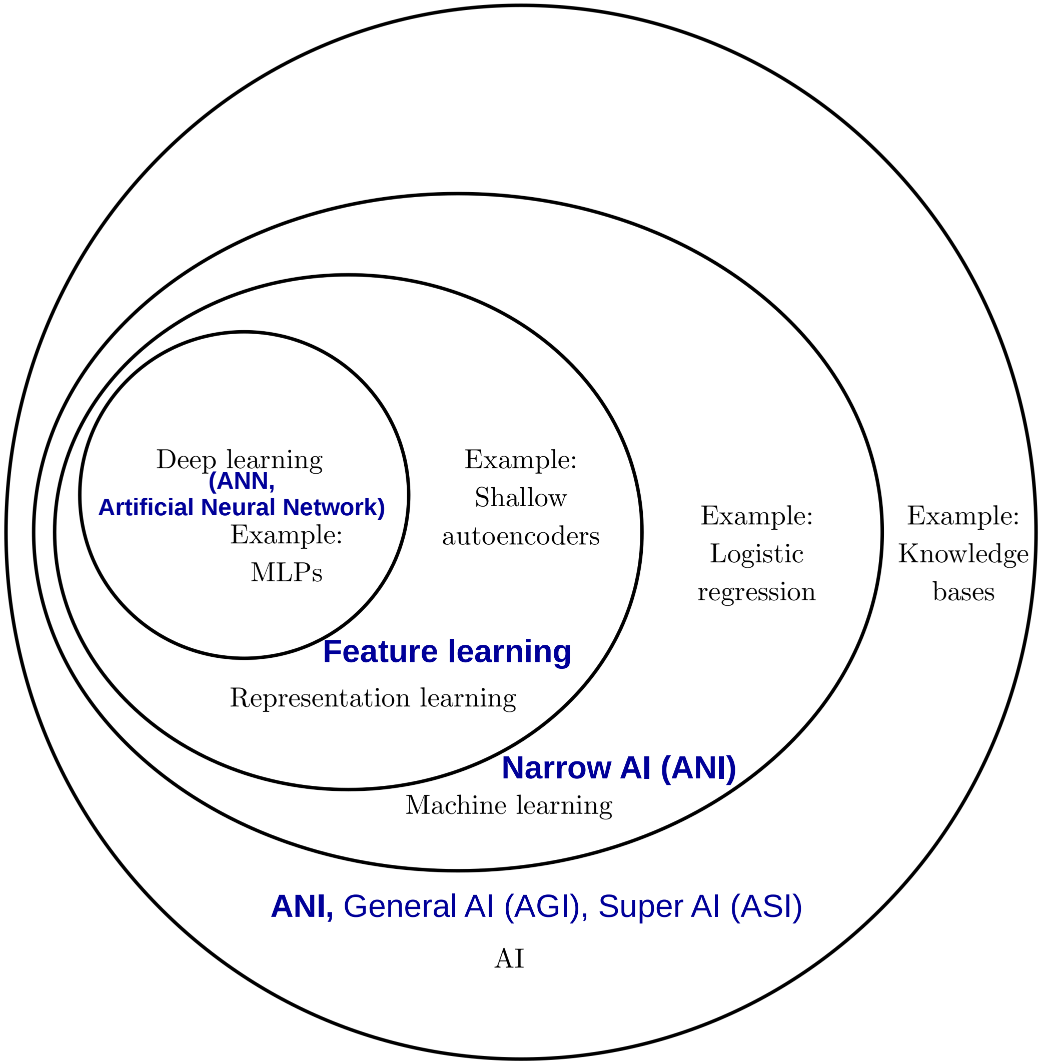
Facility/Project: HIGS
Institution: TUNL and Duke University
Country: US
Energy (MeV): 1 – 100
Accelerator: Storage Ring, 0.24 – 1.2 GeV
Laser: FEL, 1060 – 190 nm (1.17 – 6.53 eV)
Total flux: 10^7 – 3×10^{10} g/s (max ~10 MeV)
Spectral flux: 10^3 g/s/eV (max ~10 MeV)
Status: User Program
Research: Nuclear physics, Astrophysics, National Security

HIGS: an Electron-Photon Collider



Accelerator Facility
160 MeV Linac pre-injector
160 MeV – 1.2 GeV Booster injector
240 MeV – 1.2 GeV Storage ring
FELs: OK-4 (lin), OK-5 (circ)
HIGS: two-bunch, 40 – 120 mA (typ)

AI and Machine Learning



I. Goodfellow, Y. Bengio, A. Courville, "Deep Learning," 2016.
www.deeplearningbook.org

Artificial Intelligence (AI): The theory and development of computer systems able to perform tasks normally requiring human intelligence, such as visual perception, speech recognition, decision-making, and translation between languages.

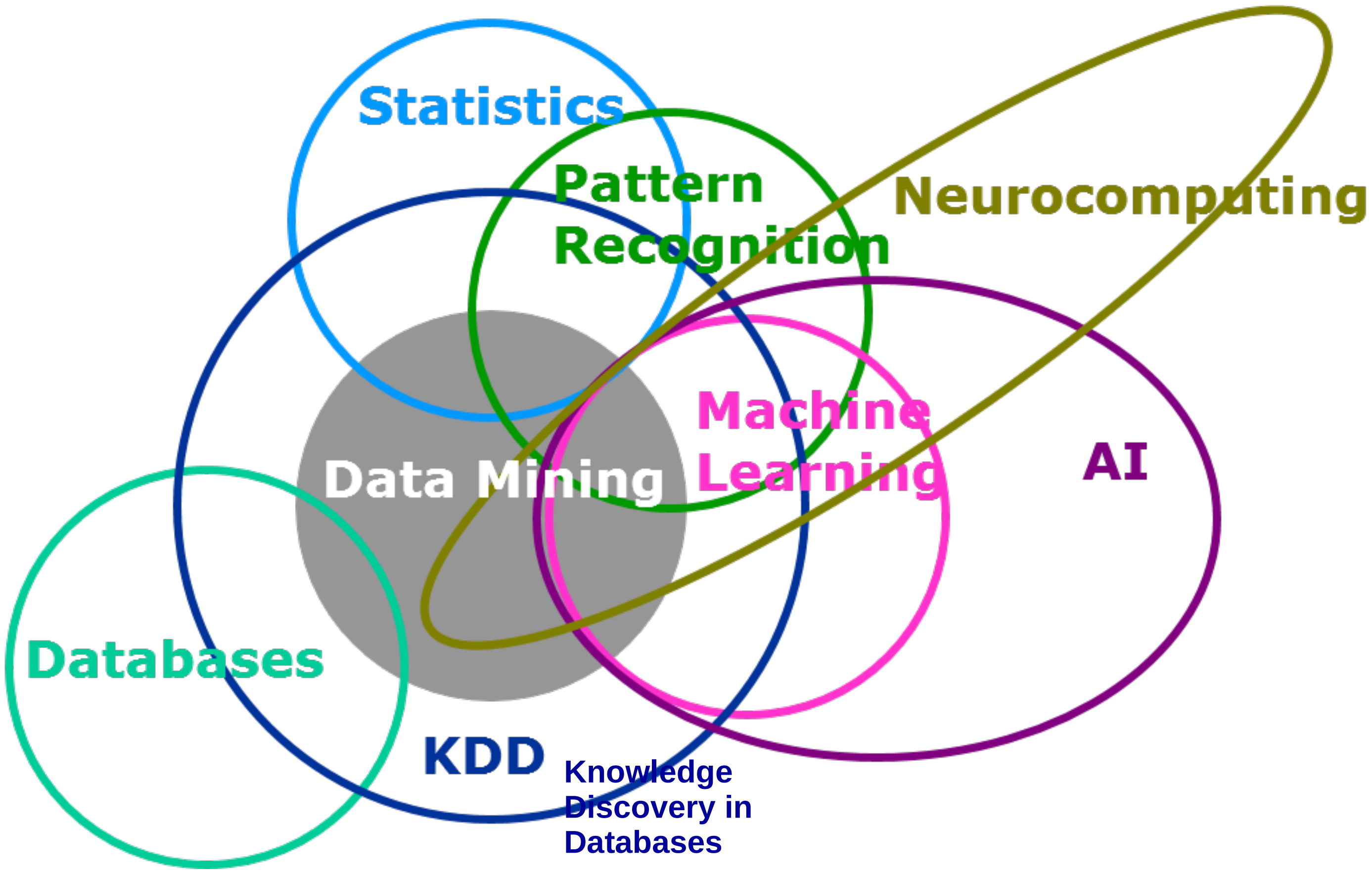
Machine learning (ML) is the scientific study of algorithms and statistical models that computer systems use to perform a specific task without using explicit instructions, relying on patterns and inference instead. ML algorithms build a mathematical model based on sample data, known as "training data", in order to make predictions or decisions without being explicitly programmed to perform the task.

ML→ Narrow AI (ANI, part of Weak AI)

- Focus on a single/limited task
- Real-time response
- Based on a specific data set
- Cannot perform outside designed task
- Examples: IBM's Watson, Siri, Google Assistant/Translate, AlphaGo

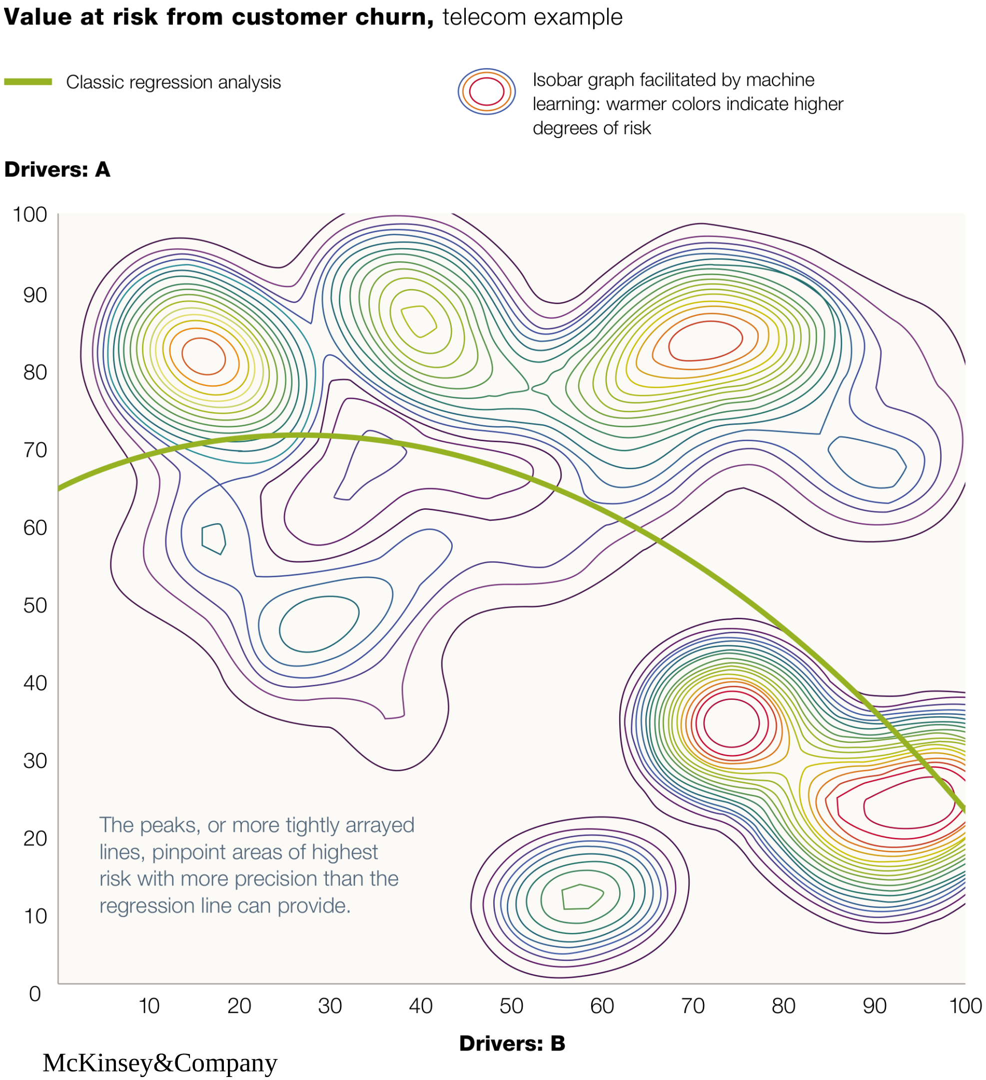
AI: The Oxford Dictionary of Phrase and Fable, 2nd Ed. (2006)
ML: https://en.wikipedia.org/wiki/Machine_learning
https://en.wikipedia.org/wiki/Weak_AI

Learning from Data—Data-driven



SAS Institute, 1998
https://blogs.sas.com/content/subconsciousmusings/2014/08/22/looking-backwards-looking-forwards-sas-data-mining-and-machine-learning/
https://www.analyticsvidhya.com/blog/2015/07/difference-machine-learning-statistical-modeling/

Contrast between statistical analysis and machine learning



Machine Learning:
An algorithm that can learn from data without relying on rules-based programming.

Statistical Modeling:
Formalization of relationships between variables in the form of mathematical equations.

McKinsey Quarterly, No. 3, 2015

AI and Machine Learning

Mathematical Models	Machine Learning
Advantages • Expert knowledge => good solutions rapidly • Input => Output with less computation • Exact solution possible • Reliable and consistent as constrained by model • First principle approach => better shield against algorithm bias	**Advantages** • “Intelligence acquisition” with refinement automated • Account for things not considered originally, but happen regularly
Disadvantages • Need expert knowledge and expert => requiring actual understanding of process and phenomenon Critical, rarely occurring details => model complexity Fine detail model => intense computation • Missing factors in the model: noise, drift, disruption, etc. • Less robust in real-world	**Disadvantages** • In reality, automation is difficulty—training requires model tweaking by an expert • Need a lot of historical data • Model training is computation intensive • Results not often predictable • Solution may not be exact, with bias from the modelers
Example: **Predict where a baseball player would hit a ball** • Build a math model using the laws of physics • Convert it to a computational model • Input: ball initial position, velocity, air resistance, etc. • Output: good prediction for where the ball will land	**Example:** **(The same baseball example)** • Build a machine learning algorithm (model) • Feed in player’s previous data: input: pitch speed, placement, etc. outcome: where the ball landed • Use the model to predict an outcome of a new input

https://www.leansystems.co/blog/machine-learning-vs-mathematical-modelling
Company founded by Sebastien Roy (U. Montreal) et al. (2016)

AI and Machine Learning

Types of ML Algorithms

- Supervised learning
- Unsupervised learning
- Reinforcement learning
- Self learning
- Feature learning
- Sparse dictionary learning
- Anomaly detection
- Association rules

ML Models

- Artificial neural networks
- Decision trees
- Support vector machines
- Regression analysis
- Bayesian networks
- Genetic algorithms

ML: https://en.wikipedia.org/wiki/Machine_learning

Areas of ML Applications for Particle Accelerators

"Opportunities in Machine Learning for Particle Accelerators,"
https://arxiv.org/abs/1811.03172v1 (2018)

- **Anomaly detection and machine protection**
- **System Modeling**
- **Virtual Instrumentation / Virtual Diagnostics**
- **Tuning, Control, and Rapid Switching Between Operating Conditions**
- **Advanced Data Analysis**

Auralee Edelen, Christopher Mayes, Daniel Bowring, Daniel Ratner, Andreas Adelmann, Rasmus Ischebeck, Jochem Snuverink, Ilya Agapov, Raimund Kammering, Jonathan Edelen, Ivan Bazarov, Gianluca Valentino, Jorg Wenninger, **Opportunities in Machine Learning for Particle Accelerators**, https://arxiv.org/abs/1811.03172v1 (2018)

ICFA beam dynamics mini-workshop: Machine learning applications for particle accelerators (SLAC, 2018)

2nd ICFA Workshop on Machine Learning for Charged Particle Accelerators (PSI, 2019)

ML-at-SLAC 1st Workshop (2019)

AI and Machine Learning

Areas of ML Applications for Particle Accelerators

“Opportunities in Machine Learning for Particle Accelerators,” https://arxiv.org/abs/1811.03172v1 (2018)

Anomaly detection and machine protection

Acc. Problem	ML Technique; Data	Outcome	Notes
• “Quench” detection: monitoring LHC **SC magnets** [1]	• Data: archived log resistive voltage data of SC magnets • **LSTM recurrent neural networks** (RNN) • Long Short-Term Memory (LSTM)→ long-range dependencies • Explore latent patterns of the data	• Predicting future voltage sequence: best RMSE=0.00104 (128 LSTM cells, 16 previous steps, batch size 2048) • Anomaly detection to be implemented using FPGA • Current system using pre-programmed triggers highly dependable • As an addition or enhancement	• Trained on common PC • Future use with FPGA (~kHz) • not used in operation
• “Quench” detection: monitoring XFEL **SC RF cavity** [2]	• Data: physics model based 2D residue data • **Support Vector Machine** (SVM) to find the 2D boundary	• Physics model based method works well: 100% accuracy with 5000 measurements • SVM is an addition and needs further improvement	• (2D linear system)
• Faulty **BPM** detection at LHC [3]	• Data: turn-by-turn BPM data • **Autoencoder**: Bad BPM has a higher loss • **Clustering**: Density-based spatial clustering	• Cannot be used alone • Used in addition to SVD analysis	

[1] M. Wielgosz *et al*. “Using LSTM recurrent neural networks for monitoring the LHC superconducting magnets,” Nucl. Instrum. Meth. A867, 40–50 (2017),
[2] A.S. Nawaz *et al*. "Self-organzied critical control for the european xfel using black box parameter identification for the quench detection system." 2016 3rd Conference on Control and Fault-Tolerant Systems (SysTol). IEEE, 2016.
[3] E. Fol, and P. Henning. “Evaluation of machine learning methods for LHC optics measurements and corrections software”. Diss. Hochschule, Eng. Econ., Karlsruhe, 2017.

AI and Machine Learning

Areas of ML Applications for Particle Accelerators

"Opportunities in Machine Learning for Particle Accelerators," https://arxiv.org/abs/1811.03172v1 (2018)

System Modeling

Acc. Problem	ML Technique; Data	Outcome	Notes
• Switch beam parameters in a THz FEL: **injector and beamline tuning** [1]	• Data: simulated data including parameters of **RF**, **quadruples**, **Twiss parameters**, **emittance**, etc. • **Reinforcement learning** with two Neural Networks	• In one iteration the controller can set up the machine to achieve close to the desired twiss parameters (initial study) • Further study to optimize emittacne and other parameters	Sim. using Superfish and PARMELA
• Predicting beam parameters: **a gun injector** at Fermilab [2]	• Data: simulated **solenoid** strengths, **gun** phases, and **cathode** images • Hybrid of a **Convolutional Neural Network** (CNN) and a fully-connected NN	• Predicting downstream twiss parameters, emittance, etc. • Mean errors between 0.4% and 3.1% of the parameter ranges	Sim. using Superfish and PARMELA; Plan to train the model using measured data.

[1] A.L. Edelen, *et al*. "Using a neural network control policy for rapid switching between beam parameters in an FEL". Proceedings of the 38th International Free Electron Laser Conference (2017)
[2] A.L. Edelen, *et al*. "First steps toward incorporating image based diagnostics into particle accelerator control systems using convolutional neural networks". arXiv preprint arXiv:1612.05662 (2016).

AI and Machine Learning

Areas of ML Applications for Particle Accelerators

“Opportunities in Machine Learning for Particle Accelerators,” https://arxiv.org/abs/1811.03172v1 (2018)

Virtual Instrumentation / Virtual Diagnostics

Acc. Problem	ML Technique; Data	Outcome	Notes
• Predicting X-ray FEL pulse properties at LCLS [1]	• Data from single-shot diagnostics (fast and slow) for electron beam and X-ray • Linear, Quadratic, **Support Vector Regression** (SVR), **ANN**	• Energy mean error below 0.3 eV (for 530 eV photon); pulse delay below 1.6 fs; spectral shape agreement at 97% • Applicable to predict for each shot of XFEL at MHz	Tested with exp. data
• Prediction of electron beam longitudinal phase space (LPS)/current profile at (1) FACET-II and (2) LCLS [2]	• Multilayer perceptron (**MLP**) regressor from scikit-learn (an open source ML library) • Data for training and validation: from simulated non-intercepting diagnostics and LPS images for FACET-II; five nondestructive measurements and LPS images measured using a transverse deflecting cavity for LCLS	• Predicting electron beam shot-to-shot 2D LPS (nondestructive) • Prediction performance depends critically on the accuracy and resolution of diagnostic inputs • Good agreement between the predicted and simulated/measured LPS profiles	Simu. and Exp.

[1] A. Sanchez-Gonzalez, *et al*. “Accurate prediction of X-ray pulse properties from a free-electron laser using machine learning”. Nature communications 8.1 (2017): 1-9.
[2] C. Emma *et al*. "Machine learning-based longitudinal phase space prediction of particle accelerators." Phys. Rev. Accl. and Beams 21, 112802 (2018).

AI and Machine Learning

Areas of ML Applications for Particle Accelerators

“Opportunities in Machine Learning for Particle Accelerators,” https://arxiv.org/abs/1811.03172v1 (2018)

Tuning, Control, and Rapid Switching Between Operating Conditions

Acc. Problem	ML Technique; Data	Outcome	Tech.
• Switch beam parameters in a THz FEL: **injector and beamline tuning** [1] (Also in System Modeling)	• **Neural networks** trained by **reinforcement learning** • Data from PARMELA simulation	• Encouraging results: with 1 iteration the controller can achieve close to the correct Twiss parameters for test beam with energies in 3–6 MeV	Simu.
• Online **undulator tapering** optimization at LCLS [2]	• **NN** with **reinforcement learning** • Other techniques: Robust Conjugate Direction Search (RCDS); Mutil-Object Genetic Algorithm (MOGA); Particle Swarm Optimization (PSO); Extreme Seeking (ES); Simulated Annealing (SA); Markov Chain Monte Carlo (MCMC)	• Optimal zig-zag taper, the pulse energy of 5.5 keV self-seeded FEL is **doubled**	Exp.
• Tuning **quad settings** of LCLS beamline [3] • Noise issue in optimization and to incorporate physics model	• **Bayesian optimization** using **Gaussian Process** Existing technique: Nelder-Mead optimization • Hyperparameters generated using historical data	Prelim results: • achieve faster optimization than hand tuning and other optimization methods • Bayesian optimization depends strongly on hyperparameters	Exp.

[1] A.L. Edelen et al. “Using a neural network control policy for rapid switching between beam parameters in an fel,” pp. 406-409,Proceedings of FEL2017.
[2] J. Wu et al. “Recent On-Line Taper Optimization on LCLS,” p. 229, Proceedings of FEL2017.
[3] M. McIntire et al. “Bayesian Optimization of FEL Performance at LCLS,”, p. 2972, Proceedings of IPAC2016.

AI and Machine Learning

Areas of ML Applications for Particle Accelerators

“Opportunities in Machine Learning for Particle Accelerators,” https://arxiv.org/abs/1811.03172v1 (2018)

Tuning, Control, and Rapid Switching Between Operating Conditions

Acc. Problem	ML Technique; Data	Outcome	Tech.
• Temperature control of **cooling water** for a normal conducting **RF gun** [1] at Fermilab • Goal: +/- 0.02 K	• Model-based predictive control (**MPC**); linearized • A rudimentary **neural network** used to relate cavity temp. to input water temp.	• Reaching +/- 0.02K control in 5 minutes vs ~23 min using the existing feedforward/PI controller	Exp. (no RF power)
• Cooling control for a **RFQ** [2] at Fermilab	• Model-based predictive control (**MPC**) • **Neural network**: Input: water temperatures at various locations, water flow rates, ambient temp. and humidity; Output: RFQ resonant frequency	• Neural network model performs well in predicting the RFQ resonant frequency due to changes in the cooling system and amount of RF heating • More training for other operation modes, and with finer granularity	Exp. (pulsed operation)
• **Ion source** control for RFT-30 cyclotron at KAERI • Highly nonlinear and complex	• **Artificial neural network** based ion source model • Generalized predictive control (**GPC**) • **Simulated annealing algorithm**	• Found a subset of ion source parameters, but already an efficient way to control and analyze the source • Will train the ion source model with diverse experimental data	Simu. using exp data

[1] A.L. Edelen et al. “Initial experimental results of a machine learning-based temperature control system for an RF gun,” p. 1217, Proceedings of IPAC2015.
[2] A. Edelen et al. “Neural Network Model Of The PXIE RFQ Cooling System and Resonant Frequency Response,” p. 4131, Proceedings of IPAC2016.
[3] Y.B. Kong et al. “Predictive ion source control using artificial neural network for rft-30 cyclotron,” NIM A806, p. 55 (2016).

AI and Machine Learning

Areas of ML Applications for Particle Accelerators

“Opportunities in Machine Learning for Particle Accelerators,” https://arxiv.org/abs/1811.03172v1 (2018)

- **Advanced Data Analysis**

Acc. Problem	ML Technique; Data	Outcome	Technology
• Measuring muon phase space volume change at MICE [1]	• Data: simulated data including the muon parameters and a **LiH absorber** • **Kernel Density Estimation** (KDE)	• Observed changes of phase space density and volume	Sim. using MAUS and G4beamline
• Study of the short electron bunch longitudinal dynamics/ microbunching instability due to the emitted THz coherent synchrotron radiation (CSR) in storage ring KARA [2]	• Clustering using the **k-means method** • Four clusters in the longitudinal bunch profiles • Each profile with a cluster label is mapped to a CSR power in time sequence • Data for training and validation: simulation based on the Vlasov-Fokker-Planck equation	• Discovery of a correlation between the electron bunch longitudinal micro-structure and the emitted CSR power • Indication of dependencies of the micro-structure properties on various machine parameters such as beam current, synchrotron frequency, and vacuum gap.	Simu.

[1] T.A. Mohayai, P. Snopok, and D. Neuffer. “A Non-Parametric Density Estimation Approach to Measuring Beam Cooling in MICE”. arXiv preprint arXiv:1806.01834 (2018).
[2] T. Boltz *et al*. "Studies of longitudinal dynamics in the micro-bunching instability using machine learning." IPAC2018, THPAK030.

Summary

Observations:

- **Many ML models and algorithms have been explored**
- **Successful ones typically involve well-defined problems and/or small systems with limited complexity**
- **As an add-on or improvement to existing techniques/methods**
- **Many have not yet been used for real-time applications**

Opportunities to Advance/Expand Machine Learning using Particle Accelerators

- **A new type of machine learning centered around complex physical systems — the accelerator, not only the data, plays a critical or even dominant role**
  - **Develop more powerful machine learning techniques by combining physics knowledge and data models**
  - **Take advantage of a rich set of realtime information from accelerator operation**