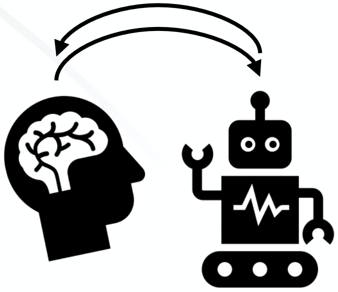


Example on how to (intelligently) augment the nuclear-data pipeline with machine learning

D. Neudecker

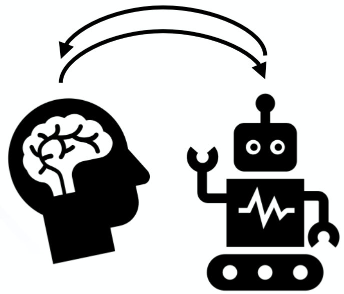
Feb. 1, 2021

Human pipeline for nuclear data, WANDA 2021

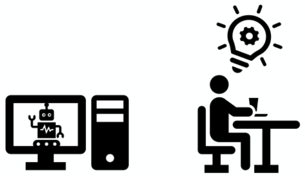
LA-UR-21-20366

The “human pipeline” vs machine learning and automation - a dichotomy??

- Machine learning helps us where human brain is overwhelmed with wealth of data. Conversely, we can integrate expert knowledge in these methods.



- Where can we use automation to free the humans in the nuclear-data pipeline from repetitive work?



- Machine learning is an exciting subject area that draws students/ Postdocs into nuclear data. (Would not say “how to train young students for this new paradigm”, but rather how can we learn along with them-***with the help of a data scientist***.) Students/ Postdocs publications in red

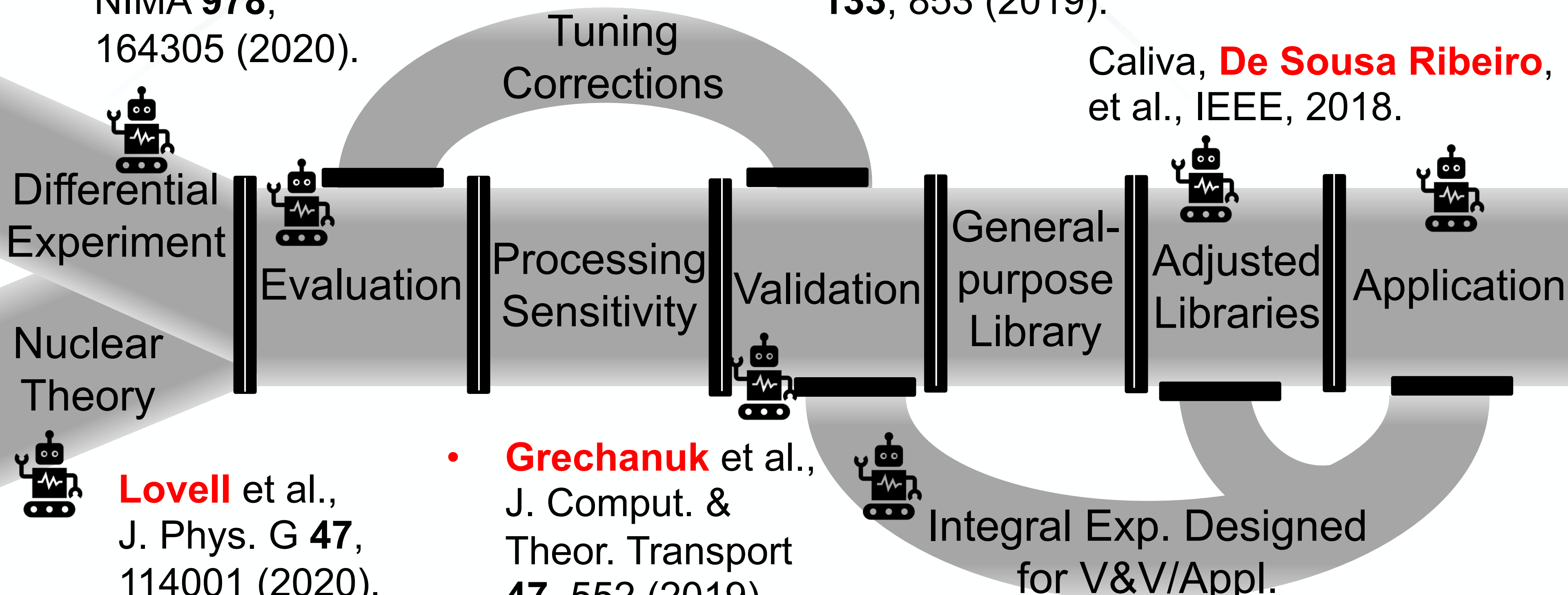
Where in the nuclear-data pipeline can automation and ML help us, a few examples:

Whewell et al., NIMA **978**, 164305 (2020).

Schnabel et al., arxiv:/2009.00521

Arthur et al., ANE **133**, 853 (2019).

Caliva, **De Sousa Ribeiro**, et al., IEEE, 2018.



Lovell et al., J. Phys. G **47**, 114001 (2020).

- **Grechanuk** et al., J. Comput. & Theor. Transport **47**, 552 (2019).
- Neudecker et al., NDS **167**, 36 (2020).

- Michaud, **Kleedtke** et al., ANS Transactions 121, 1035 (2019).
- **Siefman** et al., ANE **151** (2021).

UNCLASSIFIED

Los Alamos NATIONAL LABORATORY EST. 1943

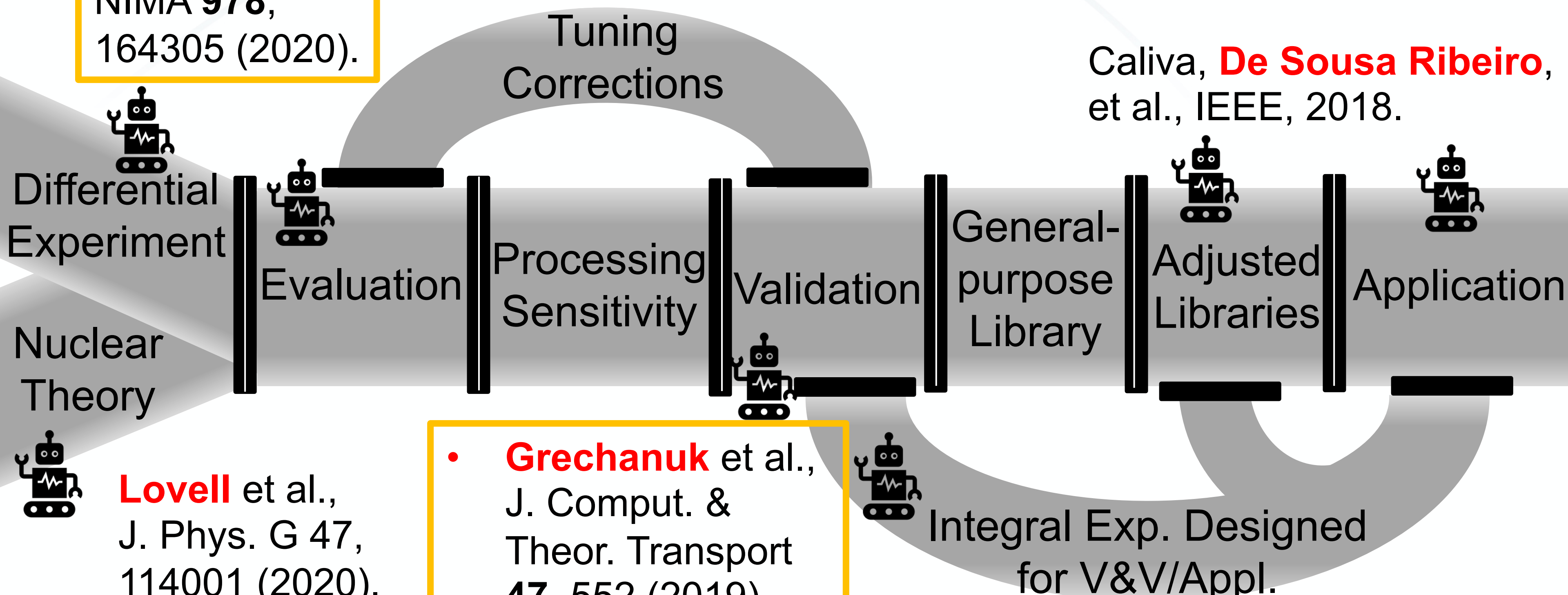
NNSA National Nuclear Security Administration

2 examples of LANL work to show how we transfer knowledge to ML and that ML augments the pipeline

Whewell et al., NIMA **978**, 164305 (2020).

Schnabel et al., arxiv:/2009.00521

Caliva, **De Sousa Ribeiro**, et al., IEEE, 2018.



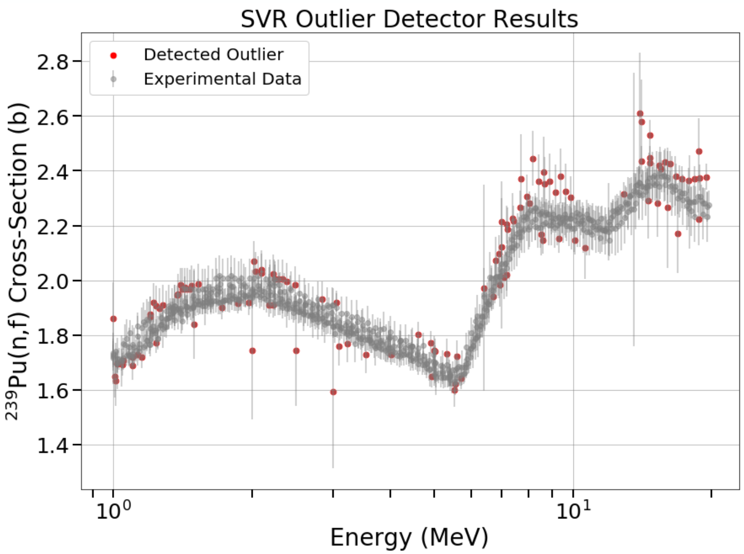
Lovell et al., J. Phys. G 47, 114001 (2020).

- **Grechanuk** et al., J. Comput. & Theor. Transport **47**, 552 (2019).
- Neudecker et al., NDS **167**, 36 (2020).

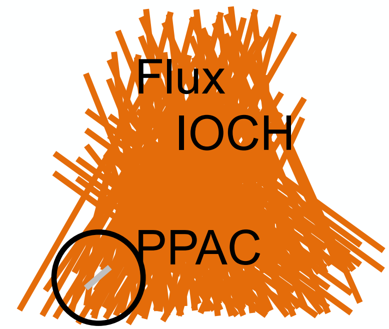
- Michaud, **Kleedtke** et al., ANS Transactions 121, 1035 (2019).
- **Siefman** et al., ANE **151** (2021).

Question: What features of differential exp. lead to systematic discrepancies/ outliers in a database?

Why traditional techniques fail: the problem has 37 feature categories (100 values) for 24 measurements. **ML can help us find trends in data, where experts are overwhelmed with data.**



Needle in the haystack …

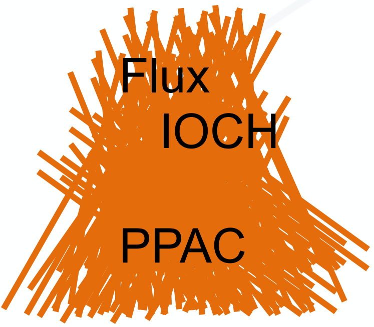


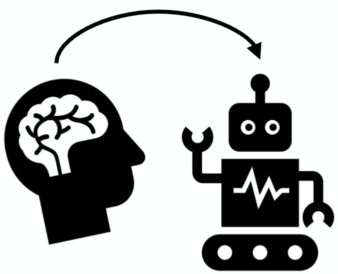
Benefit: **We are investigating the physics reasons related to discrepancies between experiments or outliers.** This can help us:

- Add missing unc/ reject data based on physics reasons → **more reliable nuclear data and uncertainties,**
- Design experiments with features known to produce reliable exp. data.

ML find physics expected features and unexpected features related to outliers, brings value to the field.

Needle in the haystack





Expert knowledge fed to ML:

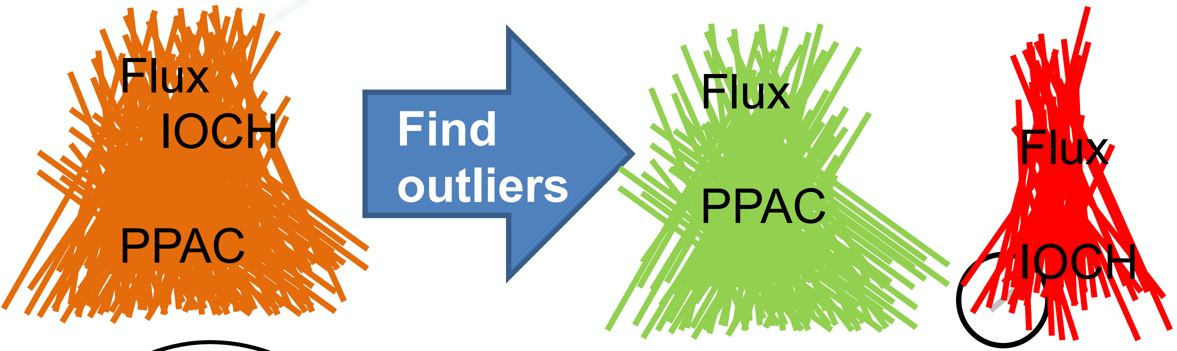
- Exp, data.
- Uncertainties
- Features

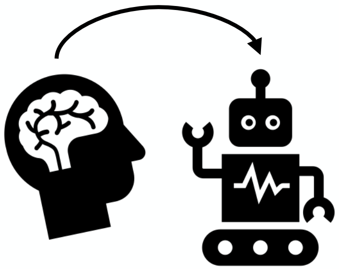
ML find physics expected features and unexpected features related to outliers, brings value to the field.

Needle in the haystack

Accepted Data

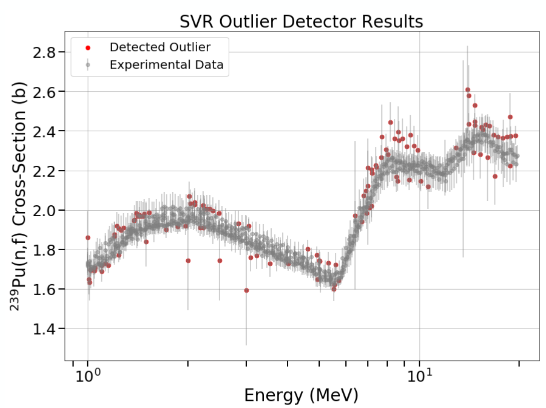
Outliers





Expert knowledge fed to ML:

- Exp, data.
- Uncertainties
- Features



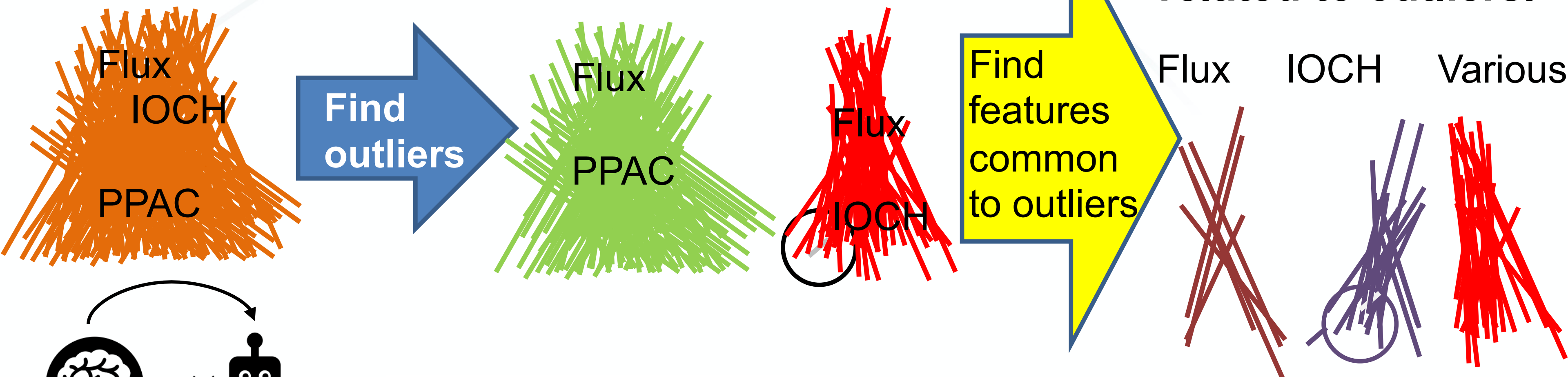
ML find physics expected features and unexpected features related to outliers, brings value to the field.

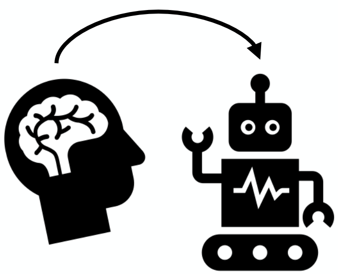
Needle in the haystack

Accepted Data

Outliers

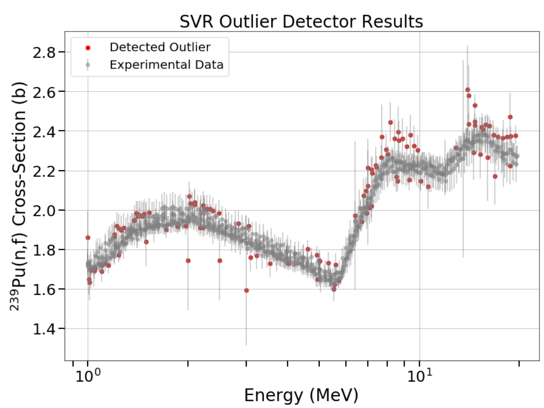
ML answer: Features related to outliers.

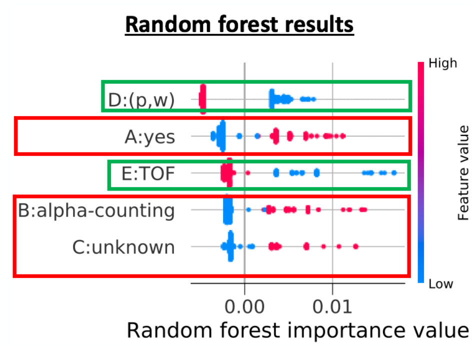




Expert knowledge fed to ML:

- Exp, data.
- Uncertainties
- Features





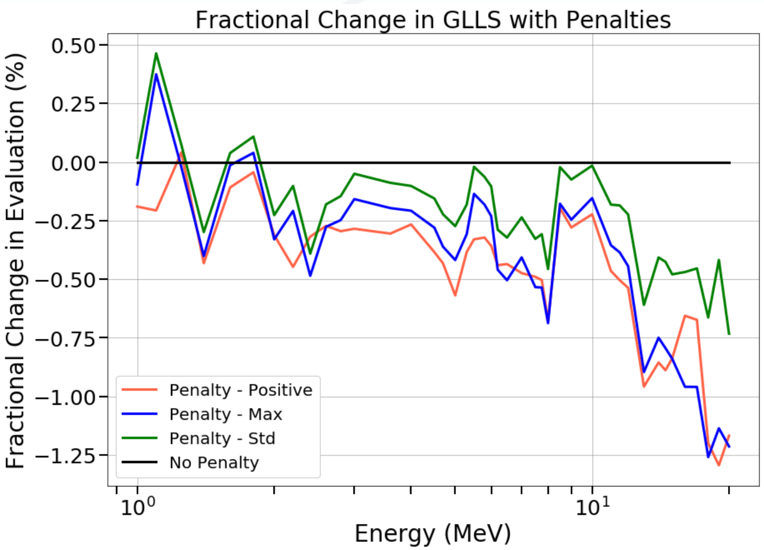
Legend

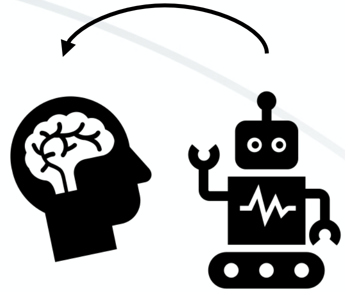
Features:
A: Normalization determined
B: Normalization determination method
C: Attenuation correction method
D: Neutron-producing reaction
E: Energy determination method

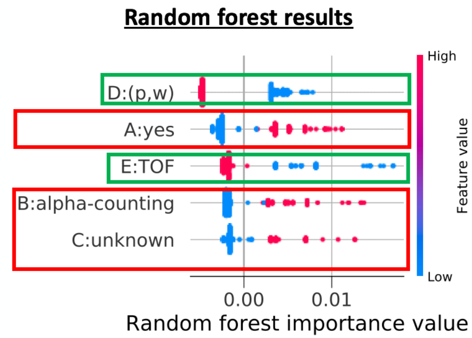
Feature value related to outlying data
Feature value related to accepted data

UNCLASSIFIED

Knowledge gained by ML can influence our evaluated data.







Legend

Features:
A: Normalization determined
B: Normalization determination method
C: Attenuation correction method
D: Neutron-producing reaction
E: Energy determination method

Feature value related to outlying data
Feature value related to accepted data

It is always up to the physicist to decide if the results are helpful. ML augments but does not replace expert judgment.

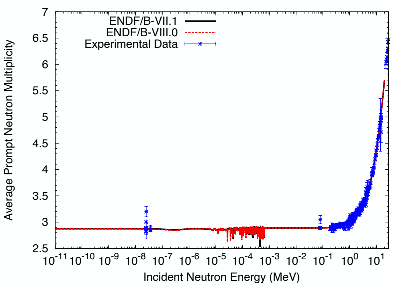
CAVEAT: for this particular problem, we lack the infrastructure to apply ML on a large scale.

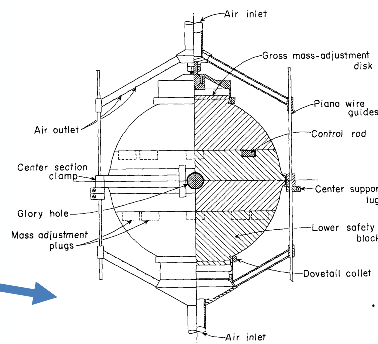
Natural language processing and SG-50 might help.

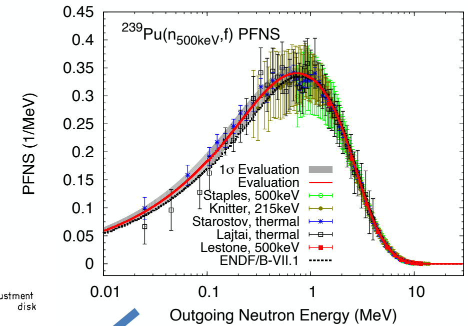
Question: What nuclear data leads to bias between simulated and experimental criticality?

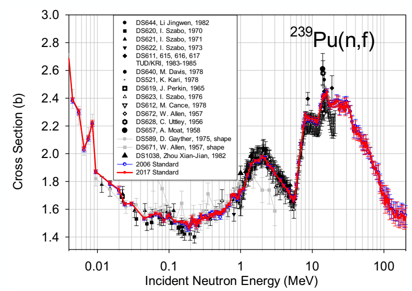
<u>Why traditional techniques fail:</u> we simulate 1 criticality value with 1000s of nuclear data. **ML can help us find trends in data, where experts are overwhelmed with data.**

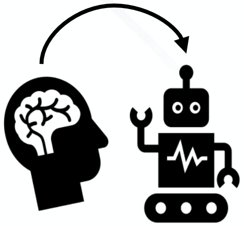
<u>Simulating the Jezebel critical assembly</u>



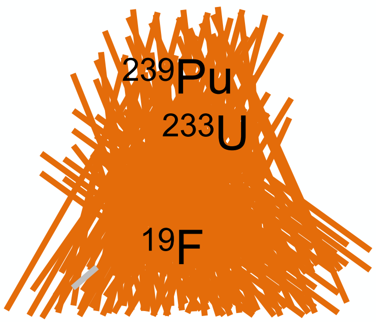








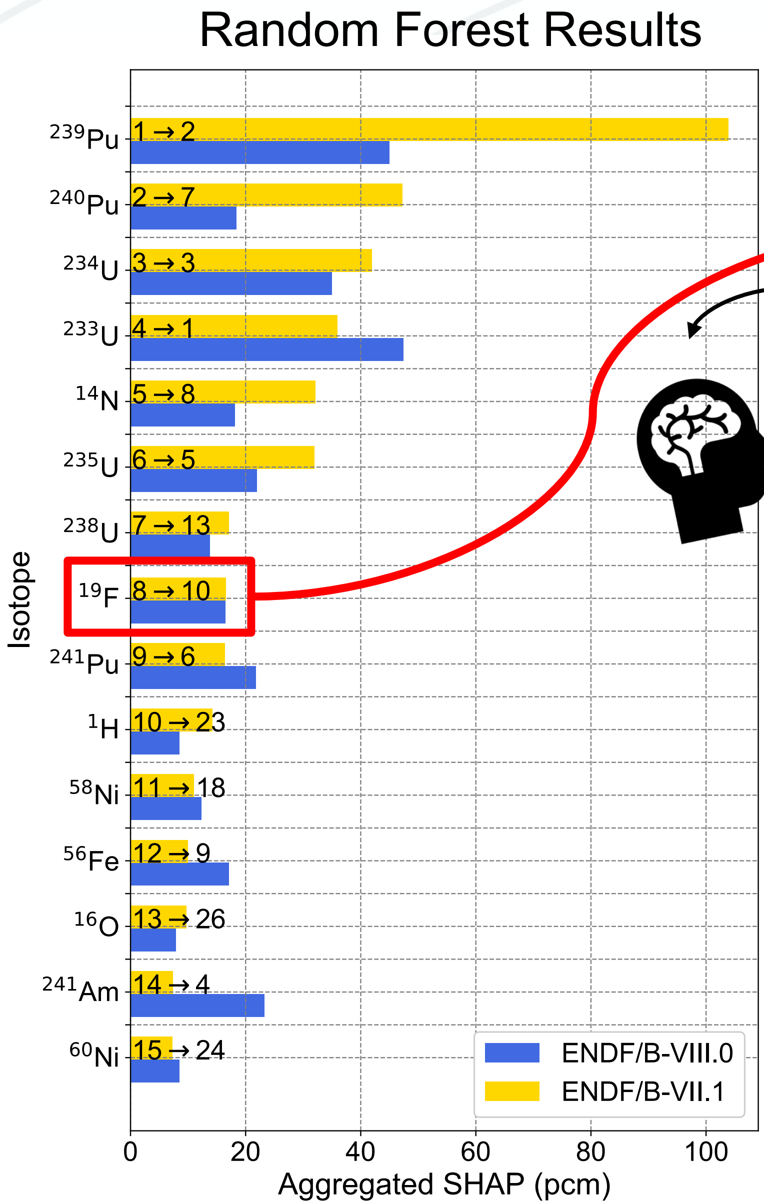
Needle in the haystack …

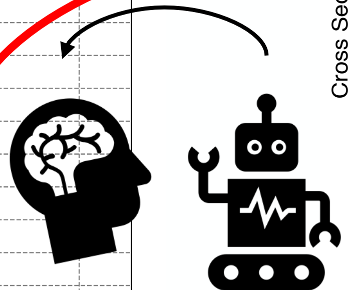


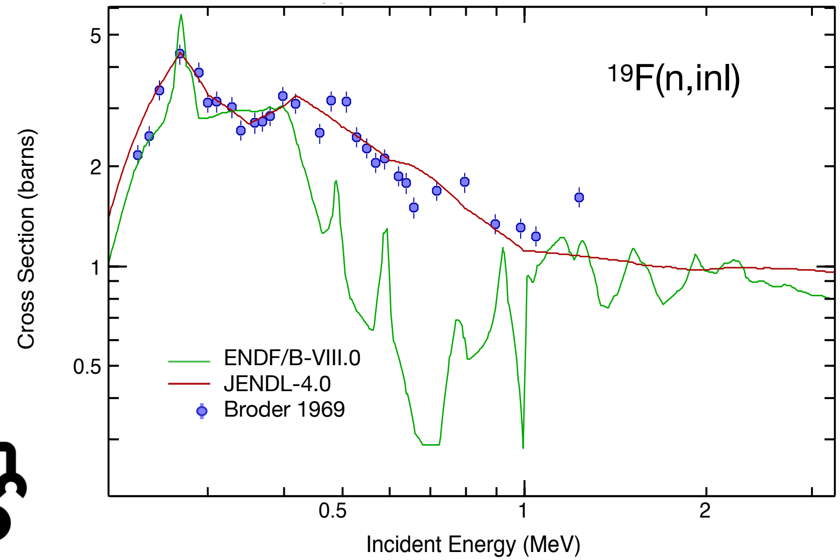
<u>Benefit:</u> **This information provides input on what nuclear data might need to be revisited and corrected:**

- Resolve timely issues in nuclear data → better data for applications,
- Identify need for integral/ differential experiments or new evaluations.

ML points towards potential issue in ^{19}F ENDF/B-VIII.0 nuclear data relevant for validation exp.







Several ^{19}F nuclear data observables, over a broad energy range, were highlighted as important to predict bias. → **Correlation effects known from traditional validation studies hamper ML because it is inherent in the data!!!!**

Conclusions:

- Can ML help our nuclear-data pipeline: **Absolutely!!!**

- ML's strength: find trends in large amounts of data where human brains are overwhelmed. This information may be crucial to improve our nuclear data.

- HOWEVER: **ML is no silver bullet. It is critical to feed it expert knowledge and use physics intuition to interpret results.** We need to:
  - Develop infrastructure and tools to provide data in an easily readable and unambiguously interpretable format (e.g., EXFOR format),
  - Develop experimental data and theory to solve physics questions,
  - Bring statisticians and nuclear-data experts together to correctly interpret the results.

Bottom line: ML is a great tool. We need to use these algorithms along with developing physics data, tools and infrastructure.

Acknowledgements

Research reported in this publication was supported by the U.S. Department of Energy LDRD program at Los Alamos National Laboratory.

We gratefully acknowledge the support of the Advanced Simulation and Computing (ASC) program at Los Alamos National Laboratory.

This work was supported in part by the DOE Nuclear Criticality Safety Program, funded and managed by the NNSA for the Department of Energy.