LA-UR-20-26936

Title:	Using Measurement Features of Experimental Data as Input for Machine Learning
Author(s):	Neudecker, Denise
Intended for:	WPEC SG-50, 2020-09-14/2020-09-15 (Paris, France) Web
Issued:	2020-09-04

Using Measurement Features of Experimental Data as Input for Machine Learning

D. Neudecker SG-50, Kick-off meeting, 9/14/20

Based on: B. Whewell, M.J. Grosskopf, D. Neudecker, NIMA, Vol. 978, p. 164305 (2020); https://doi.org/10.1016/j.nima.2020.164305

Yes, measurement features (attributes of data, e.g., detector type, sample thickness) can help us!

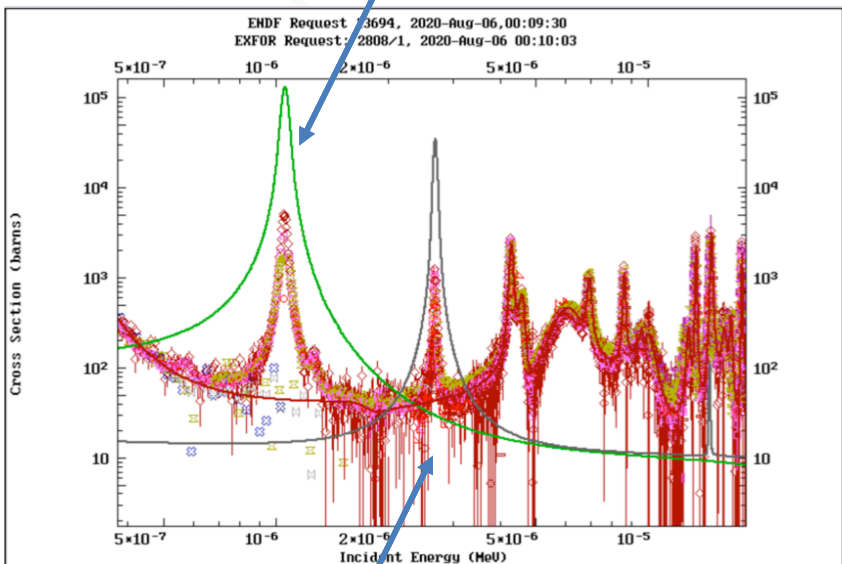
- UQ: better pin-pointing similarities between data sets. E.g., if same monitor nuclear data were used, a common unc. applies.

- UQ: Easily Retrieving unambiguously-formatted features helps in automating UQ.

- Understanding systematic effects between data sets: by comparing, e.g., contaminants across several data sets, one might be able to conclude that discrepancies are caused by them. -> ML might be needed if a lot of data and features available`.

<u>^{241}Pu(n,tot) cross section</u>

^{240}Pu contaminants



^{242}Pu contaminants

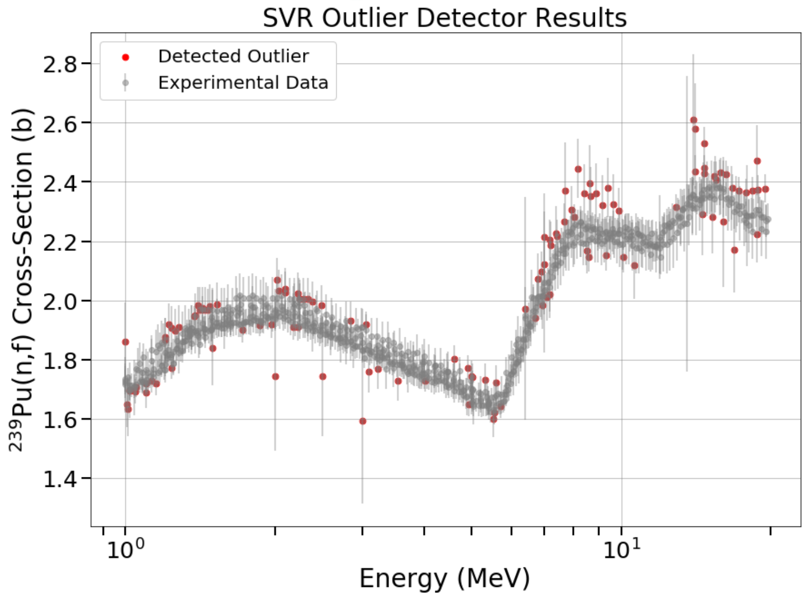
Having measurement features easily available may result in better understanding of experimental data and more reliable uncertainties -> improved evaluated data.

Example: Finding measurement features related to outlying experimental data in ^{239}Pu(n,f) cross sections with ML

Table 2: Measurement attributes that were used with this analysis. The columns labeled $x**\_$ relate the measurement attributes for the coefficient labels in the logistic regression and SHAP regression plots in Figs. 5a and 5b.

$x**\_$	Measurement Attribute	$x**\_$	Measurement Attribute
0	GMA-number	19	Neutron Flux Detector
1	Observable	20	Neutron Flux associated Particle measured
2	Absolute	21	Sample re-used
3	Ratio Isotope	22	Neutron Producing Reaction
4	Background Corrected	23	Incoming Neutron Source
5	Multiple Scattering Corrected	24	Target Backing Material
6	Attenuation Corrected	25	Target Backing Thickness
7	Stopping Power Corrected	26	Target Thickness (mg/cm^2)
8	Sample Roughness Corrected	27	Target Diameter (mm)
9	Ang. Dist. Fission Frag. Corrected	28	# atoms sample determination technique
10	Forward Boost Corrected	29	Neutron flux determination method
11	Deadtime Corrected	30	Energy Determination Method
12	Impurities Corrected	31	Background determination method
13	Random Coincidence Correction	32	Multiple scattering determination method
14	Spectrum extrapolation	33	Attenuation determination method
15	Geometry	34	Detector efficiency determination
16	Neutron flux variation	35	Impurity determination method
17	Fission Detector Type	36	Configuration of Samples
18	Reference Detector Type	37	Sample Fabrication

For each data set (24 total), 37 measurement features are given (and that are not even all we had!!!)



ML is key in finding those features that are related to outlying data points across several data sets.

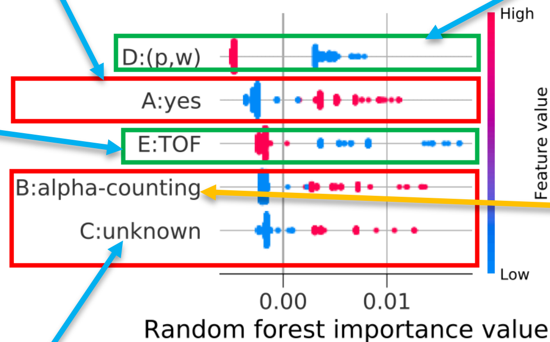
ML found features related to outliers expected by expert knowledge but also unexpected ones. ML can help evaluators!

✓ Determining normalization of (n,f) cross section poses well-known challenge

✓ (p,W) neutron source quantified well

✓ TOF is standard energy-determination technique

Random forest results



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

✗ alpha-counting is standard technique for normalization; supposedly known to 0.1% but systematic studies with the same sample across institute showed discrepancies of 0.5% → worth investigating!

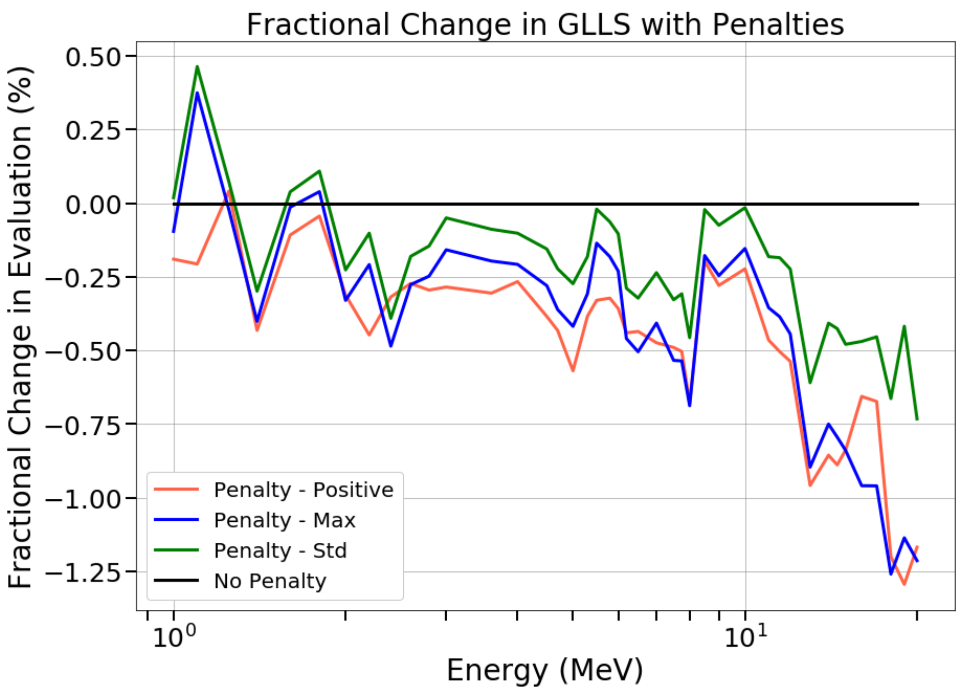
✓ Neutron attenuation expected to be poorly quantified in older data set (~60% of them measured before 1985) due to limited simulation capabilities

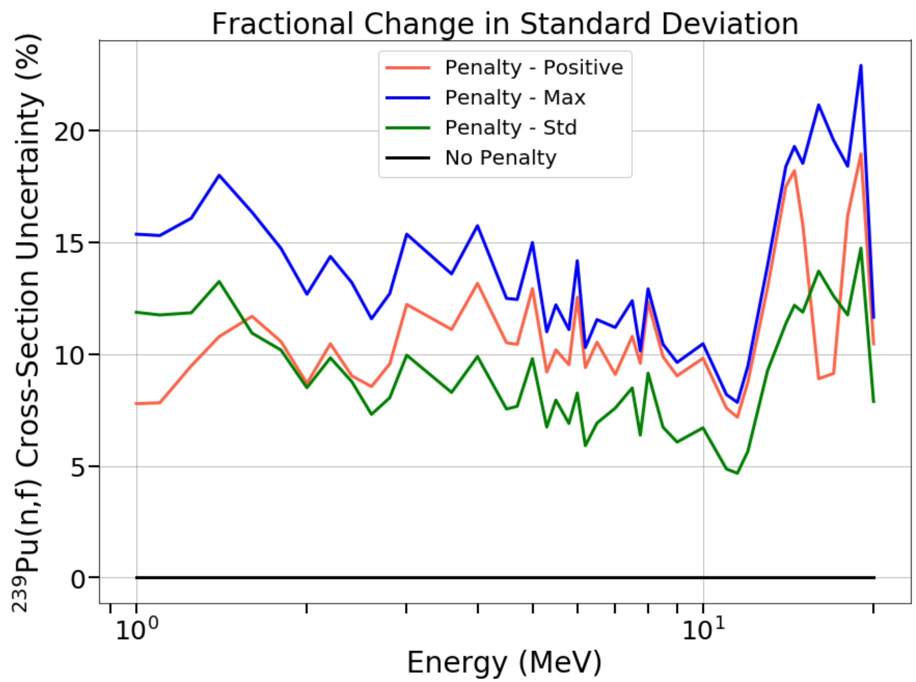
✓ **... Expected result from physics intuition**

✗ **… Surprising result**

Evaluated data and uncertainty impacted by information supplied by measurement features and analyzed by ML.

Uncertainties were added to outlying data based on the ML studies indicating that certain measurement features might be related to outlying data. -> Including these uncovered measurement uncertainties changed evaluated data and increased evaluated uncertainties.





Yes, measurement features can help us, but only if they are easy to retrieve and unambiguously interpretable!!!!!!

- It took me ~1-2 weeks to retrieve measurement features for 70 data sets:
  - Some of it can be easily copied from EXFOR (facility, neutron-producing reaction, etc.).
  - A lot of details can only be retrieved from journal articles (NOT fun!)
  - Even if information is in EXFOR, it is either ambiguously formulated (several terms often meaning the same thing, even keywords in rare cases) or in free text form.

- SG50 could help us:
  - Make this information easy to retrieve
  - Have unambiguous terms that are straight-forward to interpret.

Thank you for your attention!