ICF Design Analysis Using Machine Learning

Julia B. Nakhleh, M. Giselle Fernández-Godino, Michael J. Grosskopf, Brandon M. Wilson, and Gowri Srinivasan

Slides: LA-UR-20-23664
Narration: LA-UR-20-23712
Inertial Confinement Fusion (ICF)

• Generates nuclear fusion reactions by heating and compressing a deuterium-tritium (DT) filled capsule
  – Indirect drive: lasers heat a cavity (hohlraum) containing the DT capsule

• Plasma ignition would yield many times the input energy

• However, much of the underlying physics in ICF is not well understood
  – Better understanding of what experimental results tell us about design parameters is necessary to inform simulations and guide future experiments
Our Goal: Leverage Machine Learning to Better Understand Dominant Physical Mechanisms in ICF

• Random forest (RF) regression is an ensemble ML method that combines multiple decision trees to produce highly accurate predictions
  – Unlike neural nets, RFs don’t require large datasets to produce accurate results

• Our work uses a random forest predictor to analyze sensitivity of experimental outputs to design inputs
  – Goal: identify novel and/or unexpected input-output relationships
We can identify design parameters strongly predictive of outputs using ML interpretability metrics

- **ALE (Accumulated Local Effects)**
- **Model-agnostic global sensitivity metric**
- **Average variance in model predictions for a given range of each input feature (Apley & Zhu, 2019)**
- **Unbiased in presence of correlated features (unlike PDP)**
We will apply this to data from 140 experiments conducted at NIF beginning in 2011

- Our model uses 21 design parameters to predict 5 output parameters
  - Outputs: total yield, velocity, neutron yield, $\rho R$ from dsr, and gated X-ray BT
- Shift from high to low density hohlraum gas fills
  - Data contains ~70 high density (Group I) and ~70 low density (Group II) shots
  - We first analyze all 140 shots together, than analyze Groups I & II independently
- We used Bayesian ridge regression to estimate missing values as a function of other features (iterative imputation)
  - Poor choice of imputation method can skew model results
Correlations Between Design Inputs

- Correlated features affect importance rankings
- Removed variables: start final rise, start peak power, end pulse, dante 1 diameter, hohlraum diameter
- Rigorous assessment of parameter correlations and relationships is reserved for future work (see Fernández-Godino)
The random forest is a highly accurate predictor across outputs.
Importance analysis across inputs/outputs
Experiments can be split into high and low gas fill – We can fit the prediction model separately to capture different physical relationships
Importance rankings differ significantly between high and low density shots
Future Work

• PCA and Sparse PCA (see talk by Fernández-Godino)
  – Identify groups of correlated input variables and assess physical relationships/meaningfulness of groupings
  – Use identified groupings to inform RF and improve model performance

• Analysis of relationships with discrepancy
  – Use RF to analyze discrepancy between simulation predictions and experimental results
  – Identify the extent to which different inputs affect discrepancy
Summary

• ML provides a novel method of ICF analysis

• Random forests are able to learn and predict on experimental data with high accuracy

• Feature importance results provide insight into relationships between design inputs and measurable outputs
  – These relationships can inform future ICF design
References