# Problem Summary

The goal was to predict corium thickness in a Fukushima nuclear reactor accident scenario, where neutron emissions from **Cm-244** serve as the primary source of spontaneous fission neutrons. Mainly, if we detect a signal comes from spontaneous neutrons, we could identify the neutron source, which in turn, the location of the fuel material within the Fukushima nuclear core. Instead of direct measurement methods, which may be challenging due to high radiation fields, you used a machine learning approach to predict the corium thickness based on neutron detection data. And if we know the corium thickness, we will have information about the location of the neutron source.

This approach is necessary because:

* Corium’s composition and geometry are highly uncertain post-accident.
* Neutron interactions is complex in the melted core.
* A data-driven method can learn from simulated data and generalise to real-world scenarios.

# Data Generation & Features Selection

To create a training dataset for machine learning, Geant4 Monte Carlo simulations was used to model neutron interactions with the diamond detector in different scenarios. The main steps were:

1. Define the neutron source
	* The Cm-244 spontaneous fission source was used, as it contributes the dominant neutron emissions.
	* The neutron emission rate and energy spectrum (~2.0 MeV average energy) were based on Japan Atomic Energy Agency (JAEA) data.
2. Simulate the neutron transport
	* The corium material was modelled with different thicknesses to observe how neutron emissions change as they pass through.
	* The detector response was recorded for each scenario, including:
		+ Neutron count rate
		+ Total energy deposited (energy absorbed by the detector)
		+ Neutron source intensity (dependent on Cm-244 activity)
3. Extract dataset
	* The simulation results were automated using a Python script, which extracted and stored the key features into a CSV file.

# Machine Learning Approach

With the dataset prepared, regression models using MATLAB’s Machine Learning Toolbox was applied. The spatial distribution of Cm-244 was randomised in the

simulations to reflect real-world uncertainties in corium settling. The neutron emission rate was determined from the JAEA inventory data.

The workflow was:

1. Data Preprocessing
	* Features: Neutron count rate, neutron source intensity, total energy deposited
	* Target variable: Corium thickness
	* Validation set split
2. Model Training & Selection
	* Several regression learners were tested, including:
		+ Linear Regression
		+ Support Vector Regression (SVR)
		+ Decision Trees
		+ Ensemble Learning (Random Forest)
		+ Gaussian Process Regression (GPR)
	* GPR was selected as the best model due to:
		+ Lowest Mean Squared Error (MSE = 0.009)
		+ Lowest Root Mean Squared Error (RMSE = 0.09 cm)
		+ Ability to model complex nonlinear relationships

# Model Evaluation

To verify the accuracy of the trained GPR model:

1. Response Plot
	* Showed the predicted vs. actual values for corium thickness.
	* High alignment between predictions and actual values → confirms strong predictive power.
2. Scatter Plot
	* Compared true vs. predicted thickness values.
	* Data points clustered along the perfect prediction line → indicates high correlation and low prediction error.
3. Performance Metrics
	* RMSE: 0.09 cm
	* MSE: 0.009
	* Prediction Speed: 41,000 obs/sec
	* Optimizer: Bayesian Optimization

**In simple way:**

* Geant4 simulations provided neutron detection data for varying corium thickness.
* MATLAB’s Machine Learning Toolbox was used to compare multiple regression models.
* GPR was the best performer due to its ability to handle nonlinear neutron interactions.