Needle in a Haystack:

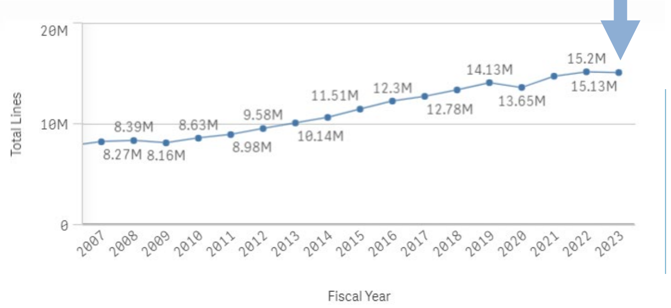
Using Machine Learning for Improved Compliance Targeting in the Human Food Program

FDA’s Center for Food Safety and Applied Nutrition, Office of Compliance

Which container would you pick?

74% increase last decade

to 15M shipment lines in 2023



Percent of commodity imported (examples)

94%
Seafood

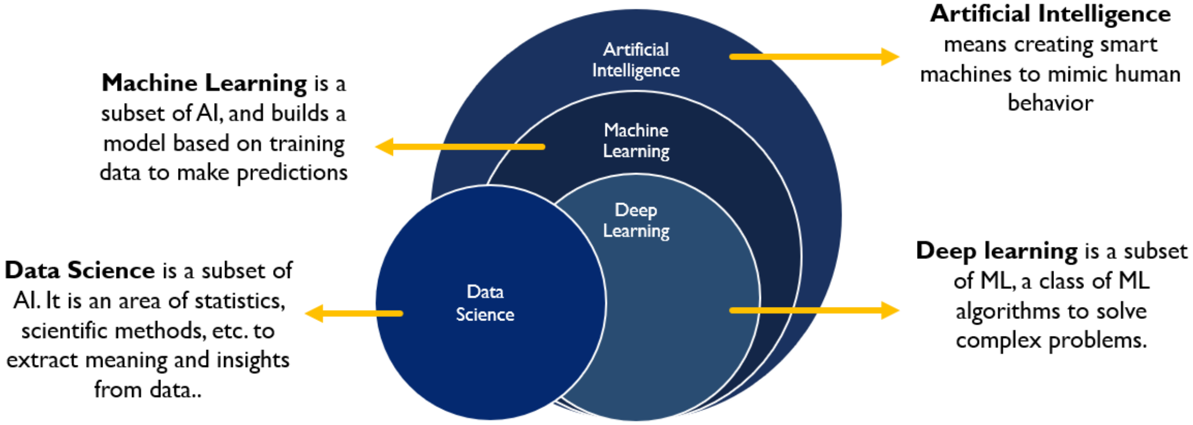
55%
Fresh Fruit

32%
Fresh Vegetables

***Source:** FDA Data Dashboard (https://datadashboard.fda.gov/ora/cd/impsummary.htm)

What is machine learning (ML)?

- The use of computational, statistical, and mathematical models to **learn patterns from historical data** and then use that to subsequently **predict an outcome for a new instance**.
- Real-life applications: Email spam detectors, credit/debit card fraud detection, etc.
- Within the Food Program, our traditional ML models are currently deployed to **enhance (not replace) risk-based targeting** of food products and supply chains likely to be violative of microbiological and chemical hazards.



While deep learning is often more accurate, that comes at a cost of complexity and explainability. Neural Nets are example here and include LLMs.

https://medium.com/@mishrasubhendu147/ai-vs-ml-vs-dl-vs-data-science-854d79bb1833

Meeting the Challenges for FDA's Food Safety Mission

- Increasing amounts of food imported and produced domestically.
- Limited regulatory resources to sample foods, inspect facilities, etc.
- Ever changing inventory and supply chains

https://www.fda.gov/food/new-era-smarter-food-safety/new-era-smarter-food-safety-blueprint

FDA's New Era of Smarter Food Safety Initiative

Goal: Expand predictive analytic capabilities via AI and ML, etc. using a progressive exploration and deployment, to include 3 pilots focusing on seafood over a 5-year period

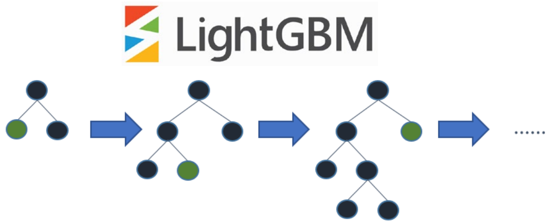
Portfolio of ML Models Developed So Far

Hazard Code	Hazard Description	Hazard Examples	Target Feature	Commodities of Interest	Domestic / Import?	Deployment Status
MIC	Microbiological (pathogenic bacteria)	*E. coli, Salmonella, Listeria*, etc.	Presence of pathogen	All Human Food	Imports and domestic	Deployed and Updated Quarterly in PREDICT
DEC	Decomposition (toxic compounds from spoilage)	Histamines, scrombotoxins, etc.	Detection of decomposition in sensory test	Seafood only	Imports only	Deployed and Updated Quarterly in PREDICT
ANT	Unapproved Antibiotics	Tetracyclines, florfenicol sulfonamides, etc.	Antibiotic concentration above safe threshold	Seafood only	Imports only	Deployed and Updated Quarterly in PREDICT
PES	Pesticides	carbendazim, glyphosate, chlorpyrifos, etc.	Element concentration above safe threshold or acceptable trace amount	All Human Food, focus on raw produce and whole grains	Imports only	Deployed and Updated Quarterly in PREDICT
ELE	Toxic elements	Lead, arsenic, mercury, etc.	Element concentration above safe threshold	All Human Food	Imports and domestic	Not Deployed - Monitoring Retrospectively
OAI	Violative Inspection (Initial or final)	N/A	Initial or final OAI classification	All Human Food	Domestic (FSMA 201 Only)	Not Deployed - Monitoring Retrospectively

Data and Modeling Process Overview

Data → **Model Training** → **Feature Evaluation**

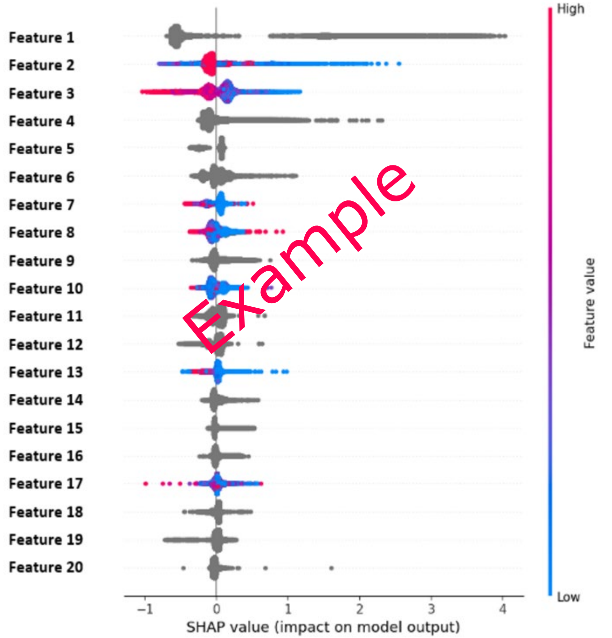
- Identify relevant features (variables) from FDA and external databases
- Clean and merge data with input from team of SME’s and data scientists



Leaf-wise tree growth

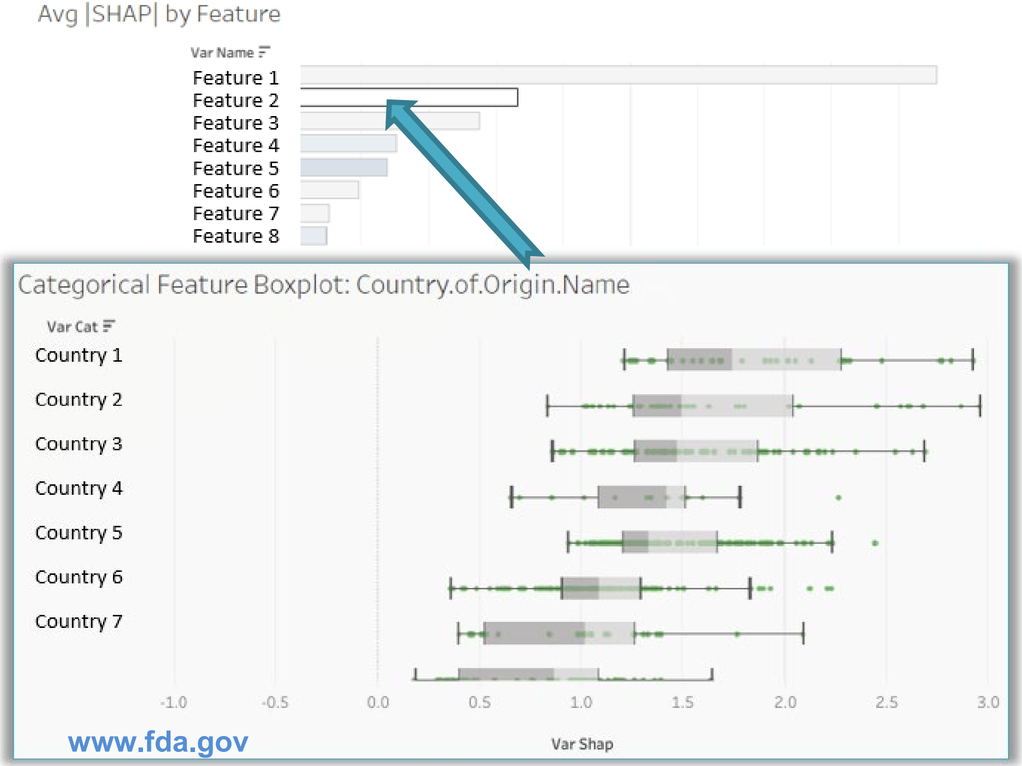
- Boosted Tree algorithm
- Target: predict violation by hazard at the supply chain level

SHAP values



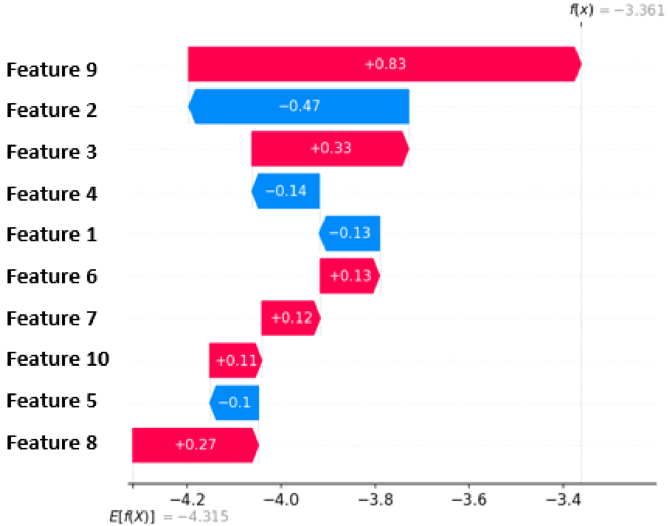
SHAP: Increasing Transparency and Stakeholder Trust in ML Modeling

SHAP Dashboard for Categorical Features (MIC-Imports model)



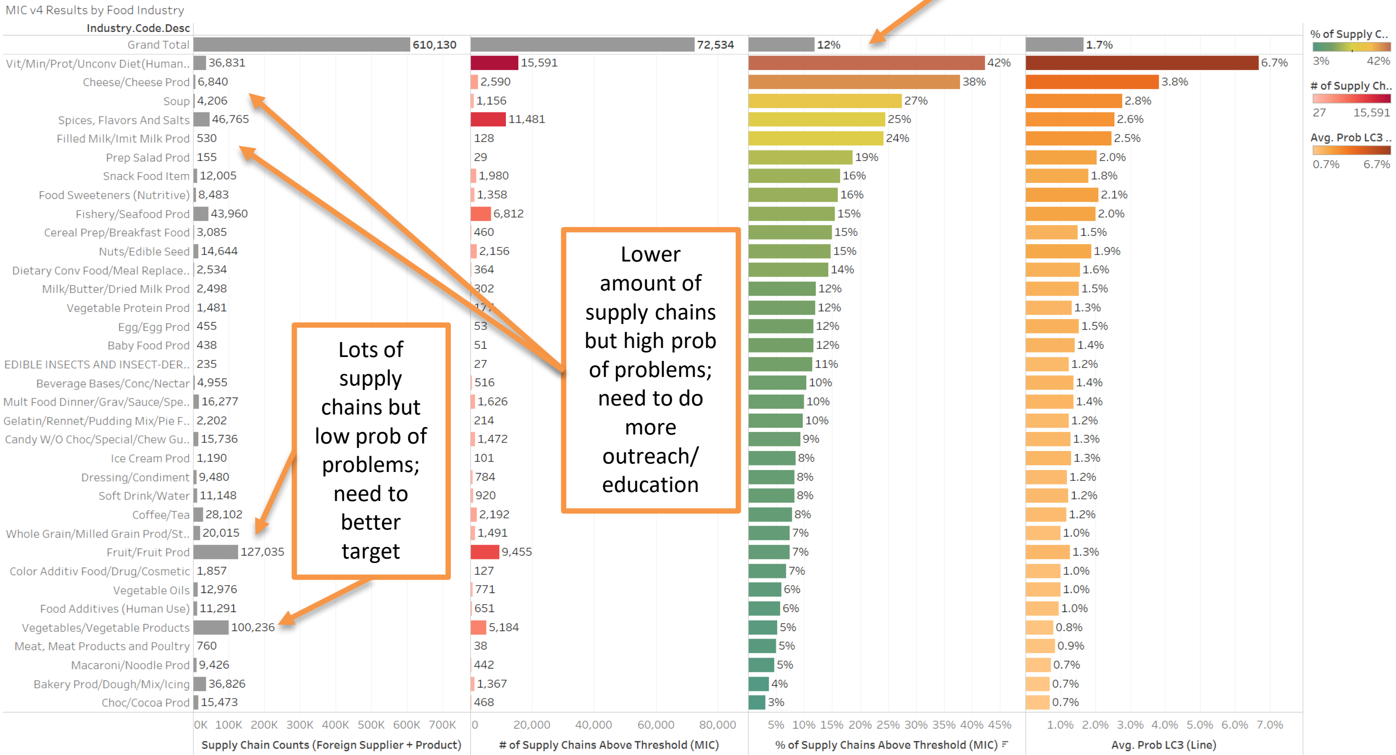
SHAP for Individual Supply Chains: Waterfall Plots

- Example: Tuna from Firm ABC
- Order of features differs by sample



Exploring ML Output by Industry

Only 12% of the active 600k supply chains are predicted violative by the model, allowing us to focus precious resources and facilitate trade

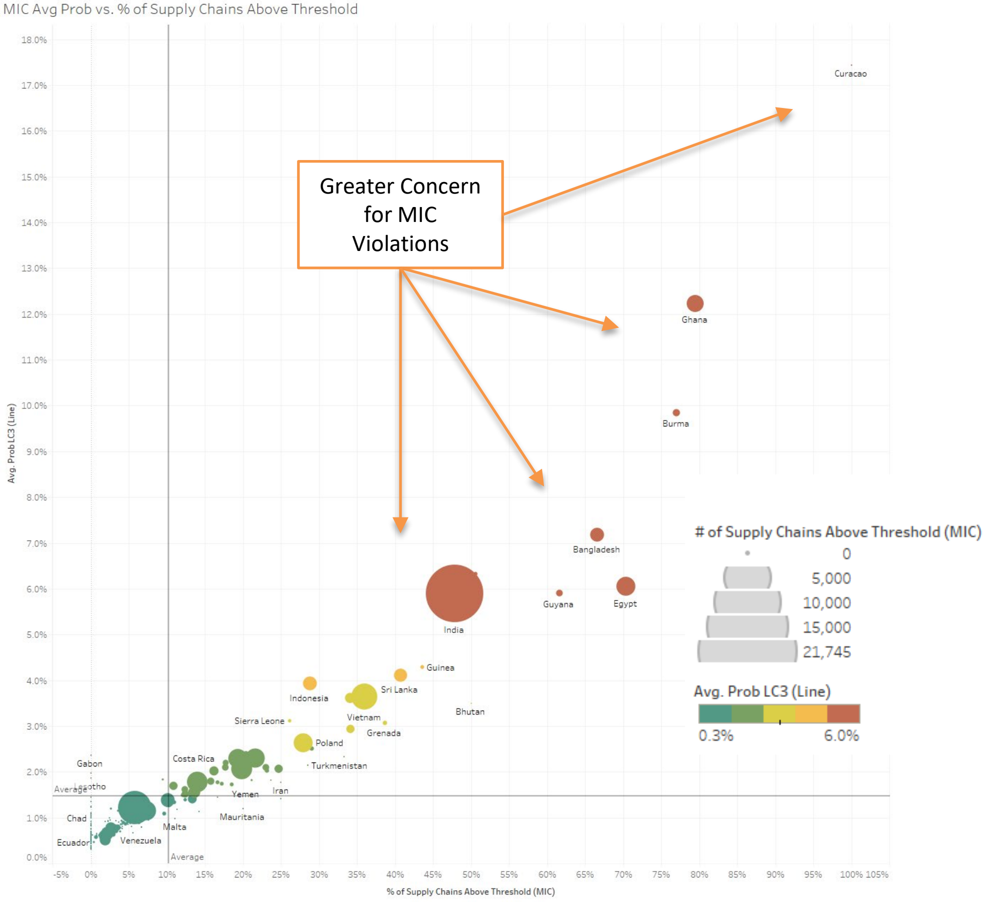


Lots of supply chains but low prob of problems; need to better target

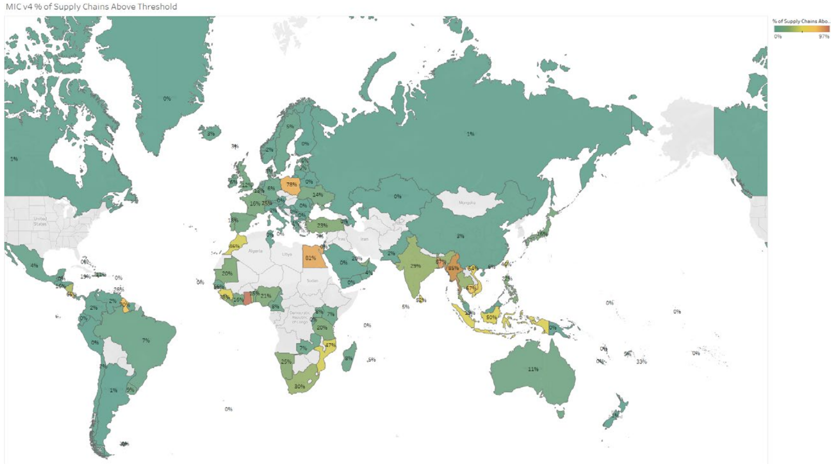
Lower amount of supply chains but high prob of problems; need to do more outreach/ education

- Produce, Spices, Seafood, Dietary Supplements, and Bakery Products had the greatest number of assessed supply chains
- The model was most concerned with dietary supplements, cheese, soup, and spices for MIC contamination

Exploring Model Output by Country of Origin (Seafood example)



- The average probability of being violative (y axis) and the % of all seafood supply chains above threshold (x-axis) vary by country
- This information may also help inform outreach efforts by country or region to improve compliance.



How we assess our models

FDA

In Silico

Train and test model on historical data; randomly assign 80/20% to train/test. Stepwise (one at a time) changes only.

Retrospective

Compare model predictions to actual accomplishment results **without influencing human sampling decisions**.

Prospective

Use model predictions **to influence human sampling decisions** and assess results compared to baseline.

Confidence

Model Results “In the Wild”

Retrospective

- Accuracy ranges from 70-92%
- Positive predictive value (PPV) is 2-5x greater than baseline, aka “hit rates”
- All models are statistically significant at 95% CI

Prospective

- Predictions shared with field staff to help inform sampling decisions (% of samples recommended by model increased)
- Results consistent or better than retrospective results (human knowledge helps)

Public Health Impact

- To date, 175 violative samples have been sampled primarily based on model recommendation
- This represents 68M KG of foods, with a declared value of $7.3M
- *Assuming* 1/2 KG serving size and 10% illness rate, **13.6M** people did not get harmed because of our ML implementation

Key Lessons Learned

1. **Data quality is essential:** Current and accurate registration, product codes, name and address of manufacturer help the model more accurately makes its predictions; missing, inconsistent, or unexpected (outliers) data are red flags.

2. **Shrinking the Haystack:** With only ~17% of active supply chains (and only 8% of total lines) predicted violative by the ML models, it greatly helps FDA focus on riskier shipments and facilities trade of the rest (win/win)

3. **Surveillance vs. Compliance:** 35% of the predicted supply chains have never been sampled ever, helping FDA address its surveillance needs while also prioritizing potential compliance violations (another win/win)

4. **Reactive vs. Proactive:** Using the ML results at the supply chain level helps us identify problem shipments and remove from the market before an outbreak or recall; at the industry/country level, it could help inform training and outreach efforts and prevent violations in the first place.

Implications of ML Results

- Better protecting public health (higher hit rate)
- Better using limited resources (focusing on subset of inventory)
- Identifying emerging trends (correcting blind spots)
- Complementing/codifying human intelligence
- Facilitating trade (not sampling those in compliance)

Moving more towards “smart” regulation

How to Move Forward Together?

FDA seeks and values input from public, academia, and industry:

- Suggestions on how to **improve models** to include additional data or features (e.g. food safety culture, weather data)
- How would external stakeholders **benefit from having access to protected, aggregated model results**?
  - Could it be used for training/outreach? Prompt preventative changes?
  - Would it improve compliance, impact purchase decisions, reduce food safety events?
- How could FDA **collaborate** with other stakeholders to share/exchange data and/or develop joint models?
  - Does industry use AI/ML to evaluate their suppliers?

Questions or Feedback?

Contact Info:

John (Chuck) Hassenplug,
Senior Policy Analyst
john.hassenplug@fda.hhs.gov