Machine Learning in Gravity Models: An Application to Agricultural Trade

Munisamy Gopinath, University of Georgia

Feras A. Batarseh, George Mason University

Jayson Beckman, USDA/ERS

NBER Conference on Agricultural Markets and Trade Policy
April 30 – May 1, 2020

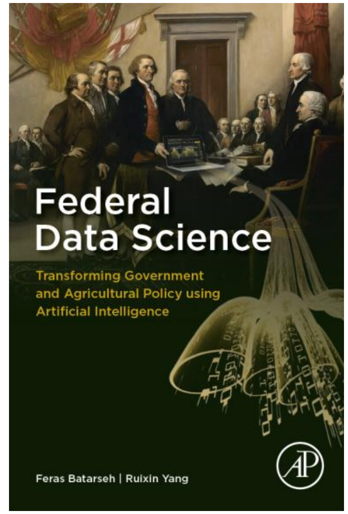
This research was supported by the intramural research program of the U.S. Department of Agriculture, Economic Research Service.

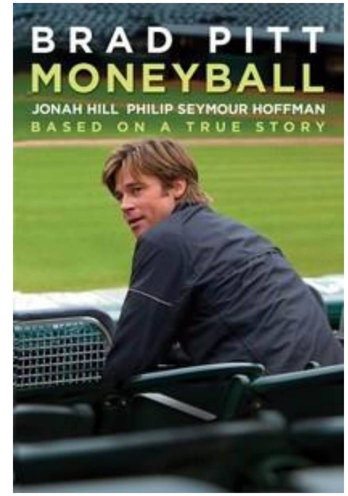
Overview

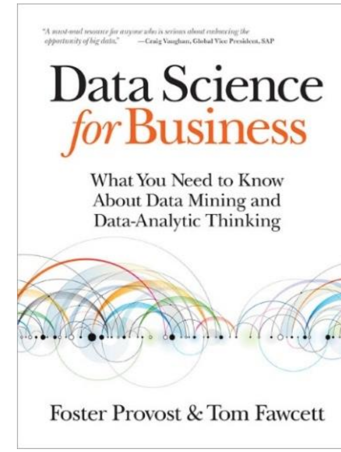
- What is Machine Learning?
- Why Agricultural Trade Patterns?
- Gravity Model and Data
- Results from Econometric and ML Approaches
- So What?

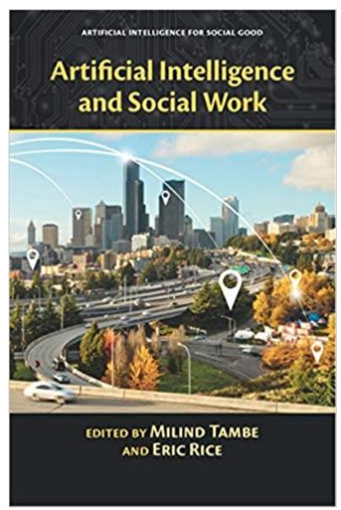
Machine Learning and AI Across domains

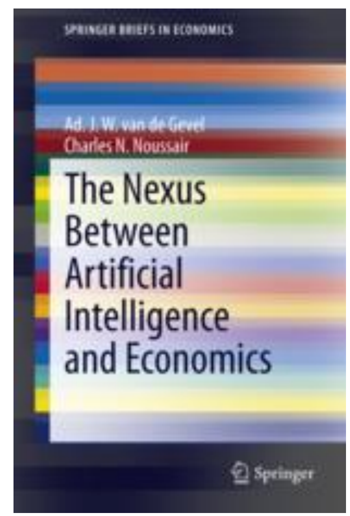
- Healthcare
- Commerce
- Energy
- Banking and Finance
- Sports
- Education
- Government & Policy

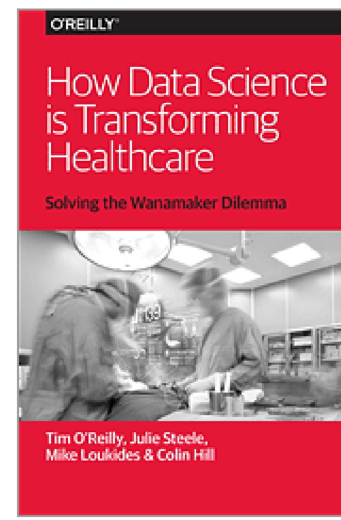












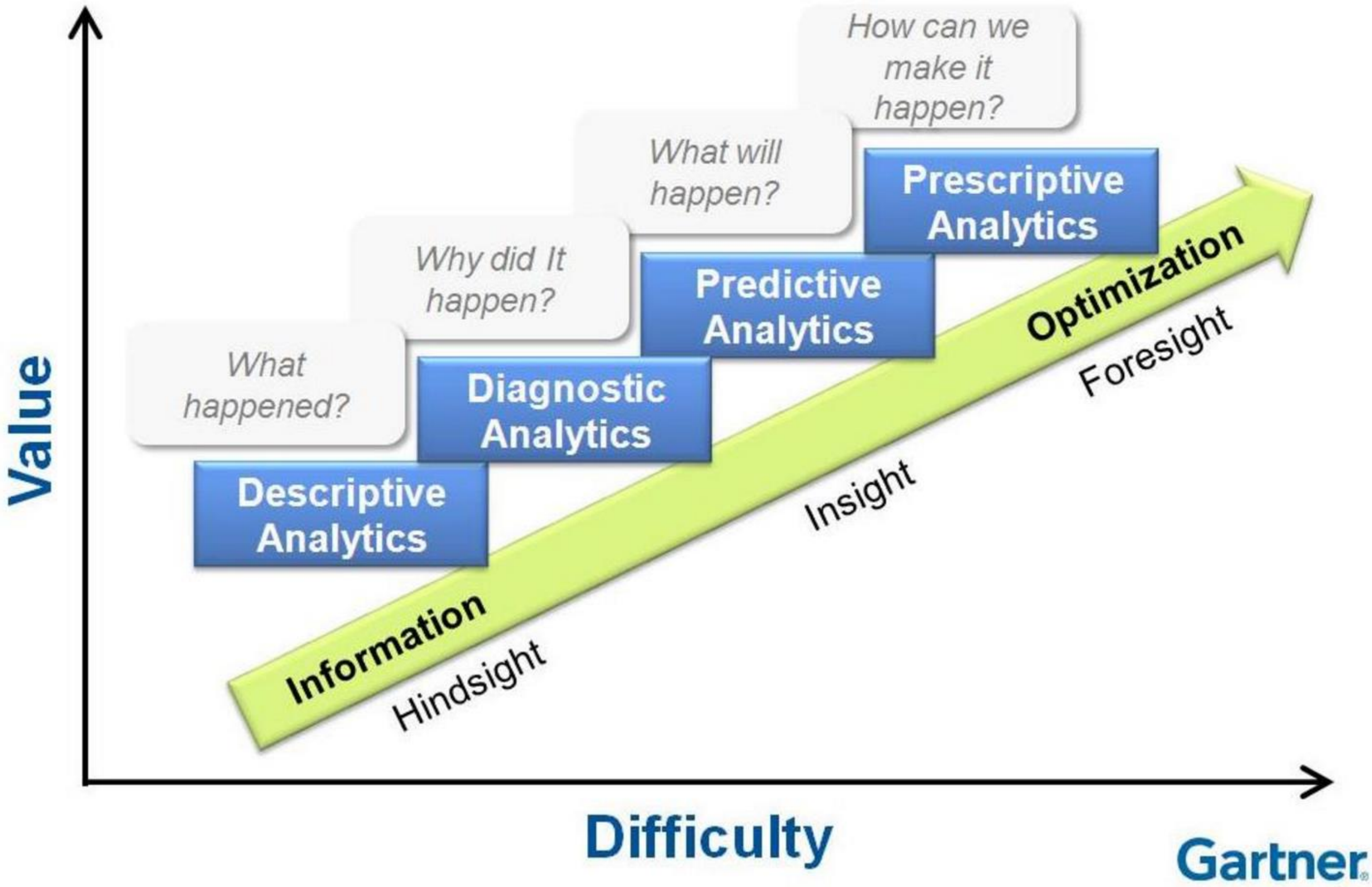
How about AI for Trade??

So, What is Machine Learning?

- A set of algorithms for advanced statistical analysis and intelligent problem-solving

- Offers a novel and flexible approach to model relationships, i.e. quantify *Y*'s response with or without a set *X* of possible predictors (supervised or unsupervised)

Four Paradigms of ML



Econometrics versus Machine Learning in the Predictive Context

- Least Squares or any other model for prediction:
  - $\left(\hat{\alpha}, \hat{\beta}\right) = \arg\min_{\alpha,\beta} \sum_{i=1}^{N} (Y_i - \alpha - \beta^T X_i)^2.$

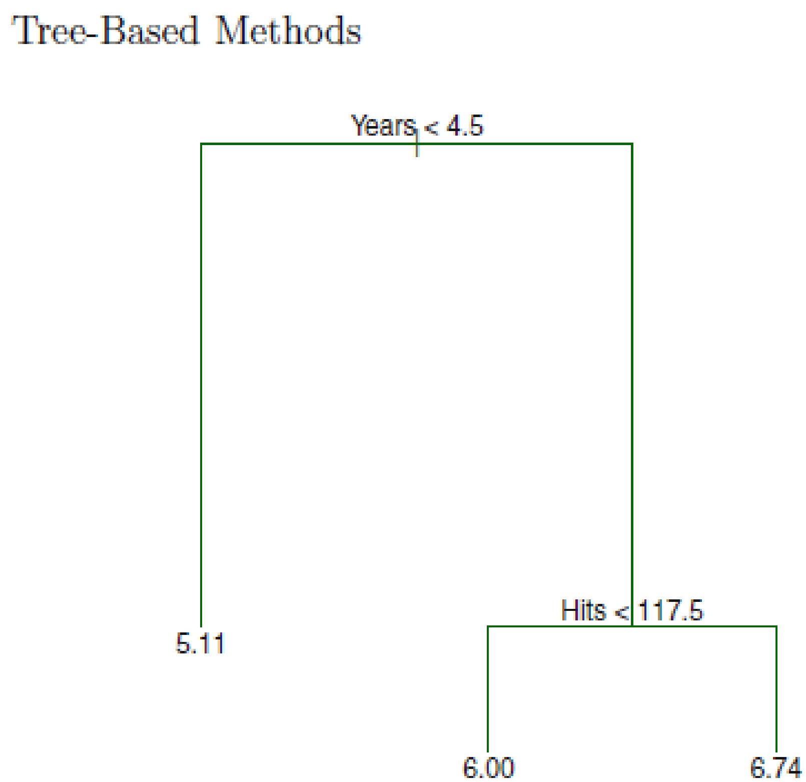
- Goal of ML, most often, is to predict Y_{N+1} from X_{N+1}. Recast that goal into a Loss function:
  - $(Y_{N+1} - \hat{Y}_{N+1})^2.$
  - Does not invoke a specific relationship between *Y* and *X*
  - Least squares is indeed an approach to minimize the loss function, but other estimators exist that dominate least squares

What Other Approaches? Machine Learning?

- Regression, auto-regressive moving average, and other forecasting models
  - for predictions and time series analysis
- Decision trees, random forests, and multiple classification algorithms
  - for decision making and categorizations
- Bagging, boosting and stacking
  - for improving weak learners, and tuning the outputs
- Clustering, associations, and correlation analysis
  - unsupervised outputs and pattern recognition models
- Neural networks, deep learning and other ensemble ML methods
  - advanced bio inspired models

- New techniques emerge every month!

Decision Trees: An Example
Baseball Salaries by Experience and Performance



Decision Tree – The Math

- Total-sample sum of squared errors for outcome *Y* is given by:

$$Q = \sum_{i=1}^{N} (Y_i - \bar{Y})^2 \qquad\qquad \bar{Y} = \sum_{i=1}^{N} Y_i.$$

- After a split based on one of the predictors (X_k) using the threshold $X_k < c$, the sum of total-sample squared errors is:

$$Q(k,c) = \sum_{i=X_{ik} \leq c} (Y_i - \bar{Y}_{k,c,l})^2 + \sum_{i=X_{ik} > c} (Y_i - \bar{Y}_{k,c,r})^2,$$

  - where *l* and *r* denote left and right of X_k using the cut-off *c* and
  - $\bar{Y}_{k,c,l} = \frac{\sum_{i=X_{ik} \leq c} Y_i}{\sum_{i=X_{ik} \leq c} 1}, \qquad \bar{Y}_{1,c,r} = \frac{\sum_{i=X_{ik} > c} Y_i}{\sum_{i=X_{ik} > c} 1}$.

Why Agricultural Trade?

The recent trade wars have challenged economists in predicting trade flows (patterns) across countries.

- Agricultural trade has been caught in the recent tariff crossfire.
- Agricultural trade reforms have been a sensitive issue (e.g. Doha, TPP)
- Trade policy uncertainty opens up the possibility that alternative approaches may be needed to make better forecasts.

We rely on the popular gravity model, but employ ML tools to answer three questions:

- Which economic variables (such as GDP and population) are likely associated with a country's exports?
- Can ML algorithms ensure learning and explain predictions from *country-commodity-year* cubical trade data?
- Can ML techniques qualitatively improve the forecast relative to that from traditional econometrics or applied/computable GE models?

Gravity Model

- Applied to the standard gravity specification
  - $Y_{ijt} = g(X_{it}, X_{jt}, i, j, t)$
  - Y_{ijt} is bilateral trade between country i and country j at time t
  - $X_{it(jt)}$ is the set of possible predictors from both countries
  - Set $\{i, j, t\}$ refers to a variety of controls on all three dimensions

Why Primary Commodities?

- Wheat
- Beef
- Corn
- Soy
- Sugar
- Milk
- Rice






- Data availability (country, commodity and time)

Data

- Bilateral trade data - Global Agricultural Trade System (GATS)
- U.S. ITC's Gravity Portal (2019) for gravity variables
  - Over 70 variables, but 35 chosen based on correlation (cardinality in ML)
- MacMap – Tariffs (since 1988 only)

- Data available for 1960-2017/18, but vary across commodities

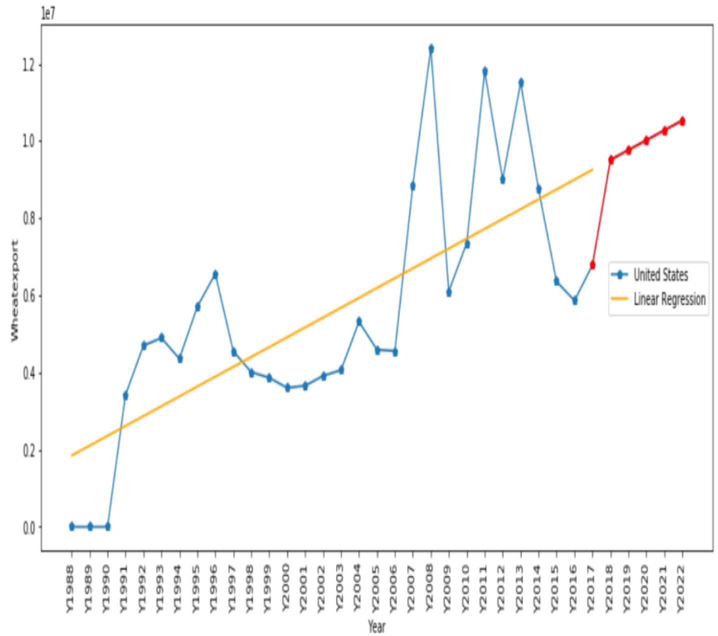
Traditional Approach

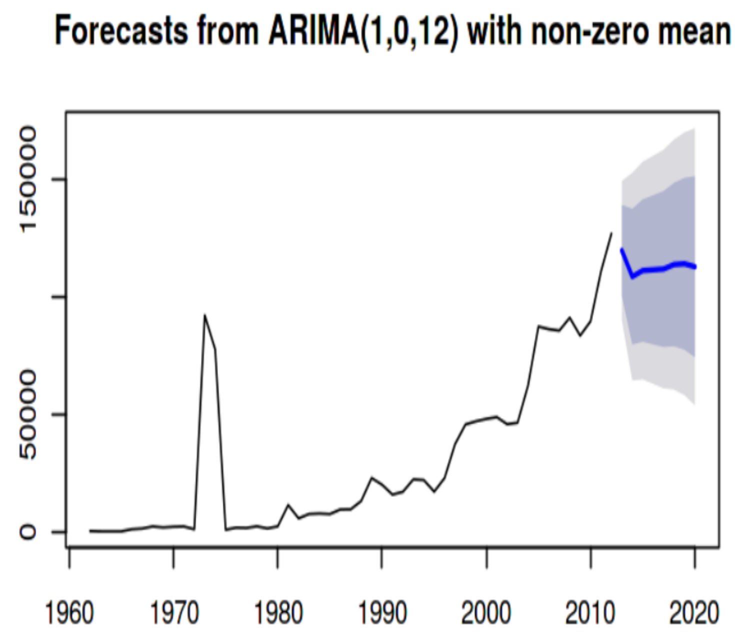
- Poisson Pseudo-Maximum Likelihood Approach
  - Recent method with all the bells and whistles
    - Zero trade
    - Heteroskedasticity
    - Exporter-time or importer-time dummies for multilateral resistance
  - PPML rarely used for prediction
  - Has been challenged in quantifying economic significance

ML Technologies used

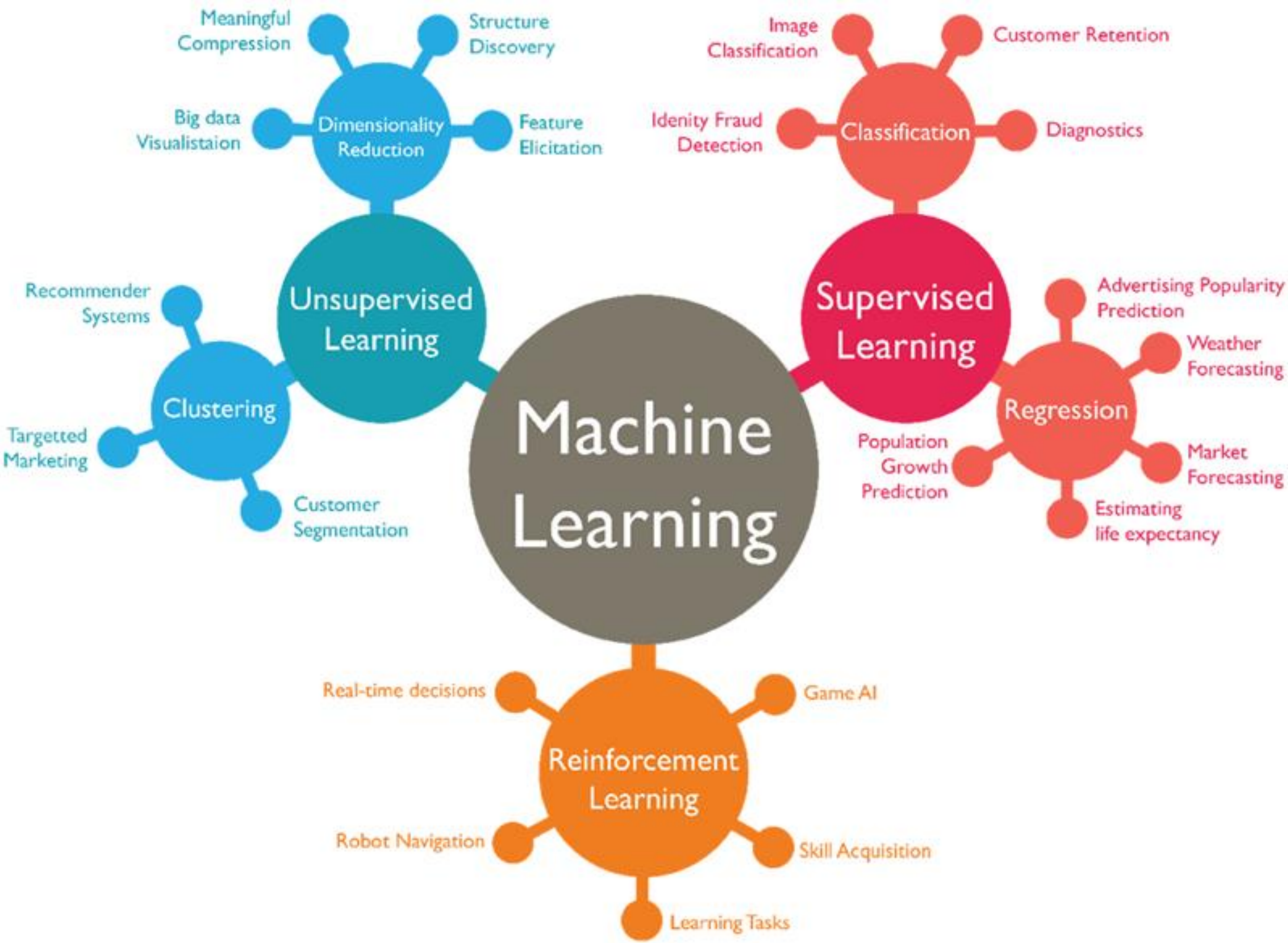
- Neural Networks
- Boosting
- Bagging
- ARIMA
- Linear Regression

Regression and ARIMA Predictions (included in the ML Toolbox)





Supervised and Unsupervised Methods



Boosting and Bagging in Supervised Methods

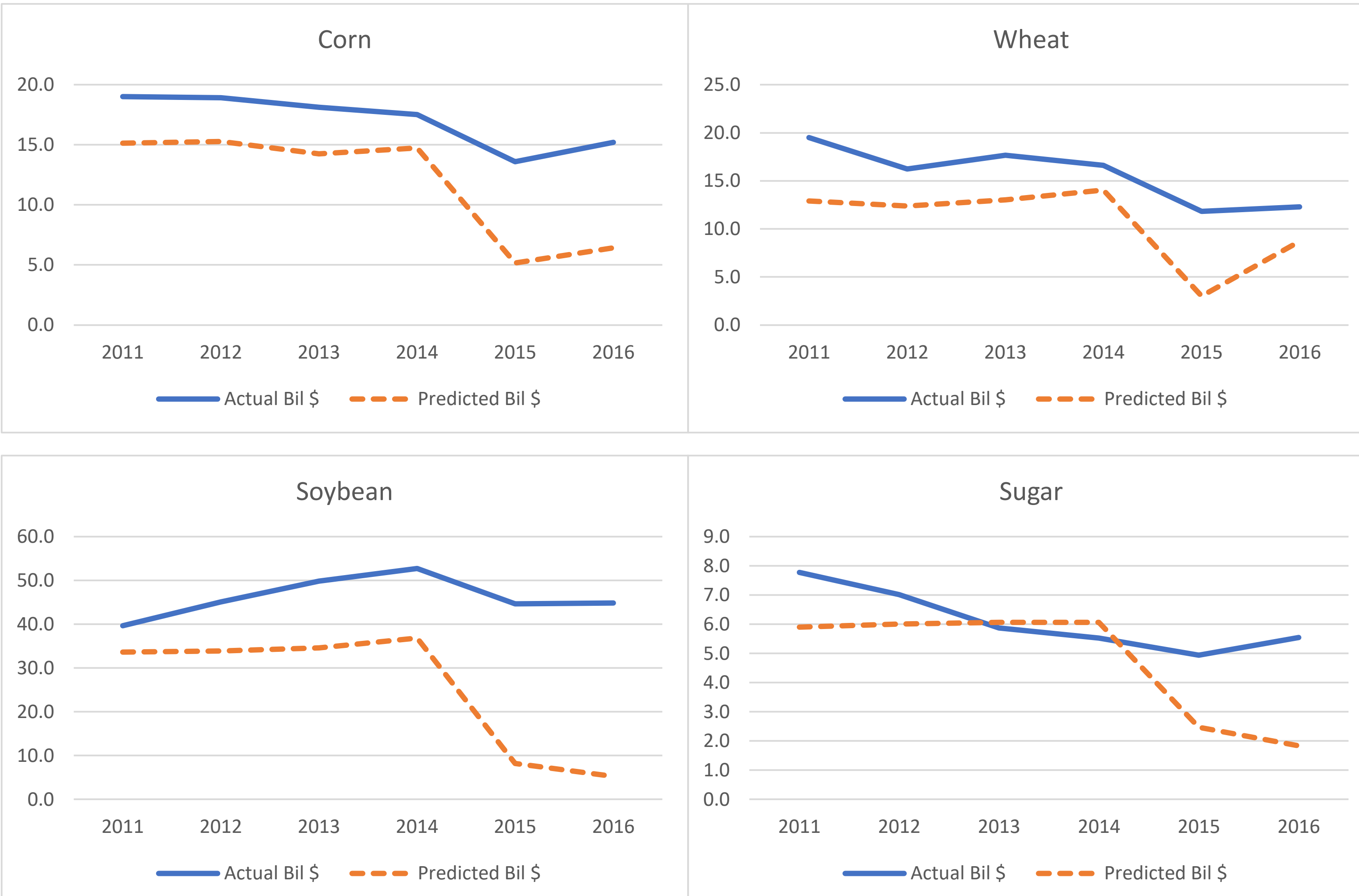
Similarities		Differences
Both are ensemble methods to get N learners from 1 learner…		… but, while they are built independently for Bagging, Boosting tries to add new models that do well where previous models fail.
Both generate several training data sets by random sampling…		… but only Boosting determines weights for the data to tip the scales in favor of the most difficult cases.
Both make the final decision by averaging the N learners (or taking the majority of them)…		… but it is an equally weighted average for Bagging and a weighted average for Boosting, more weight to those with better performance on training data.
Both are good at reducing variance and provide higher stability…		… but only Boosting tries to reduce bias. On the other hand, Bagging may solve the over-fitting problem, while Boosting can increase it.

Supervised ML Model Results

Commodity	Observations	LightGBM	XGBoost	Random Forest	Extra Trees Regression
Beef	Training – 27153 Test – 7290	0.538	0.598	0.560	**0.601**
Corn	Training – 29500 Test – 10583	0.680	0.624	**0.723**	0.678
Milk Powder	Training – 58434 Test – 15594	0.782	0.772	0.787	**0.828**
Rice	Training - 47697 Test – 12750	0.426	0.416	**0.451**	0.423
Soybean	Training – 22448 Test – 6018	0.593	0.616	**0.649**	0.581
Sugar	Training – 28660 Test – 7644	**0.448**	0.347	0.439	0.447
Wheat	Training – 26520 Test – 7212	**0.670**	0.498	0.643	0.665

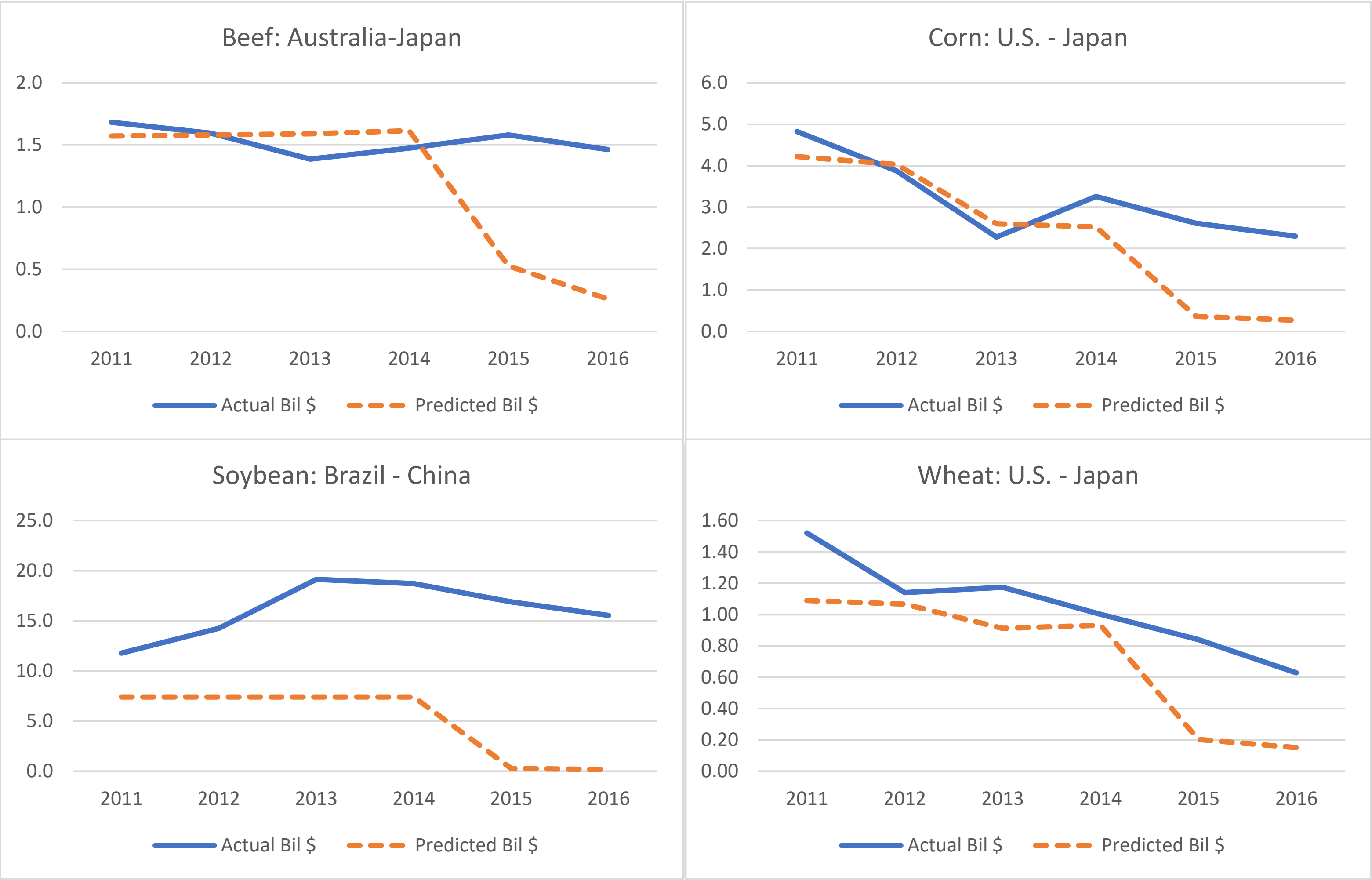
SUPERVISED

Total Trade Projections – Extra-trees



Bilateral Trade – Extra-trees

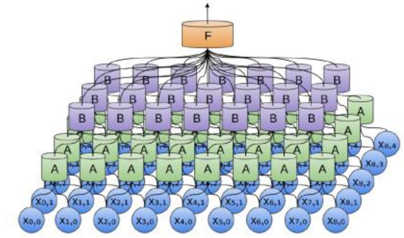
SUPERVISED



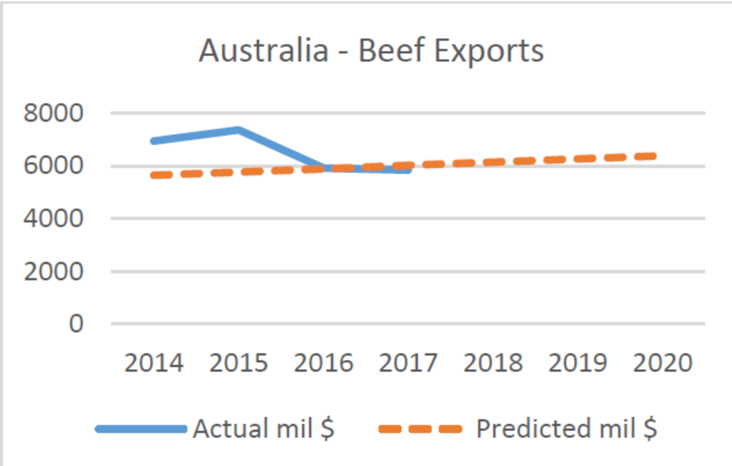
Economic Significance – Supervised Model

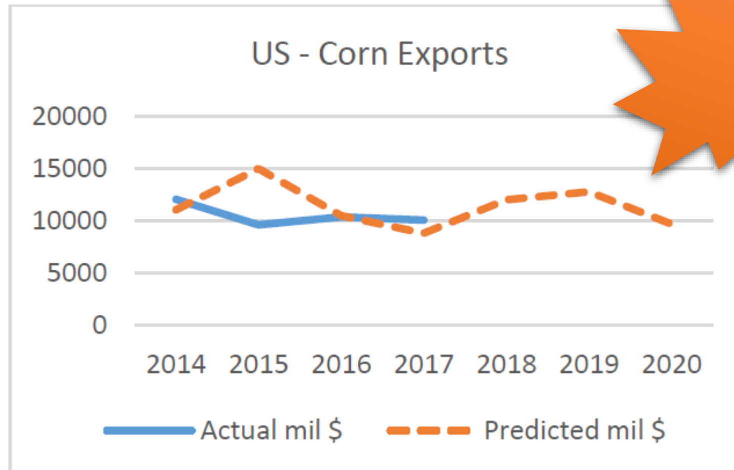
Variables	Beef	Corn	Milk Powder	Rice	Soybean	Sugar	Wheat
Population_Origin	100	100	100	100	67	43	100
Population_Destination	78	9	48	80	84	100	30
Distance	9	8	61	69	89	70	33
GDP Per Capita_Origin	55	90	34	63	30	18	40
Longitude_Destination	66	15	20	26	100	39	44
Latitude_Destination	53	16	54	41	14	92	17
Longitude_Origin	34	12	81	35	39	9	48
Latitude_Origin	92	39	34	49	5	17	72
GDP Per Capita_Destination	27	6	36	31	30	29	27
Time	10	5	53	33	55	19	17
Tariffs	15	2	13	52	6	24	9

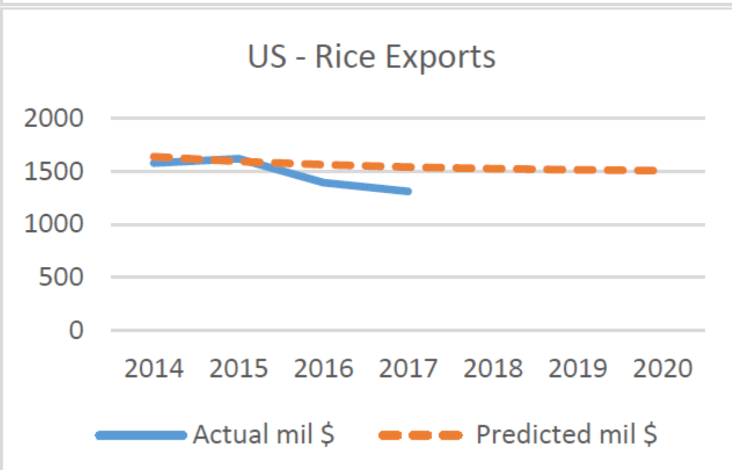
Unsupervised Model Predictions

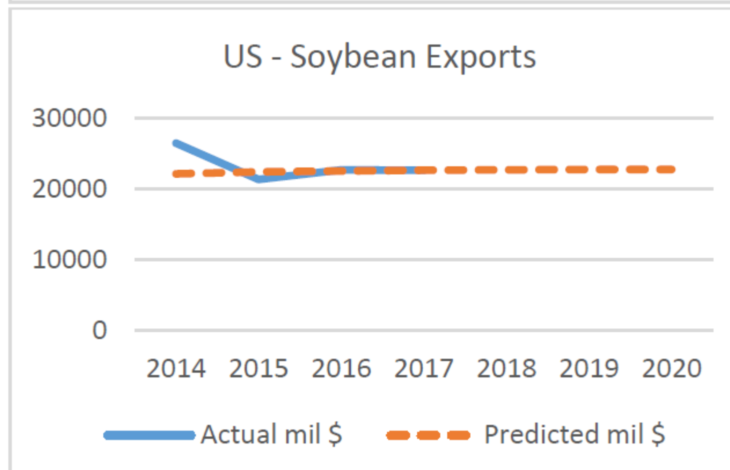


Using DL

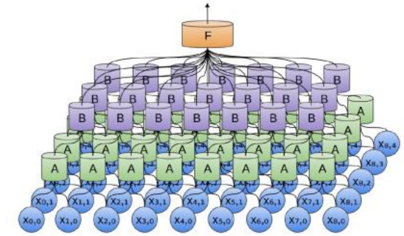


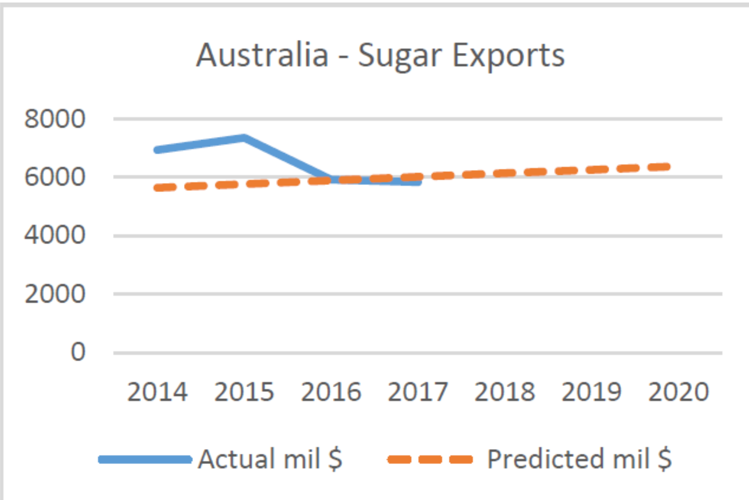


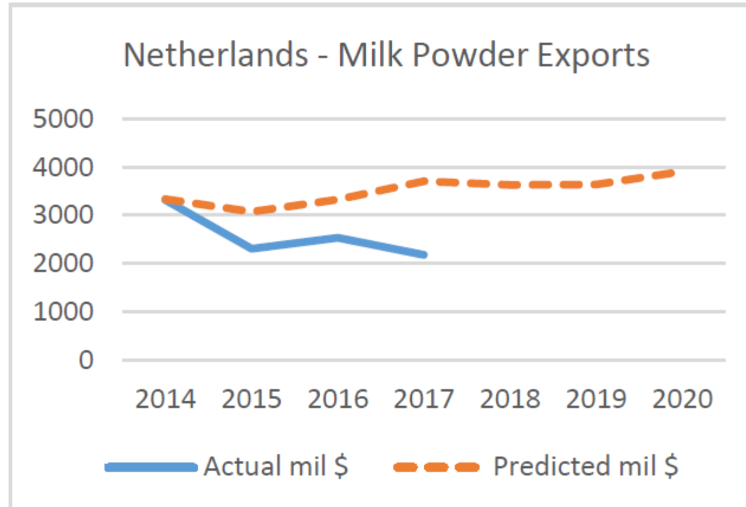




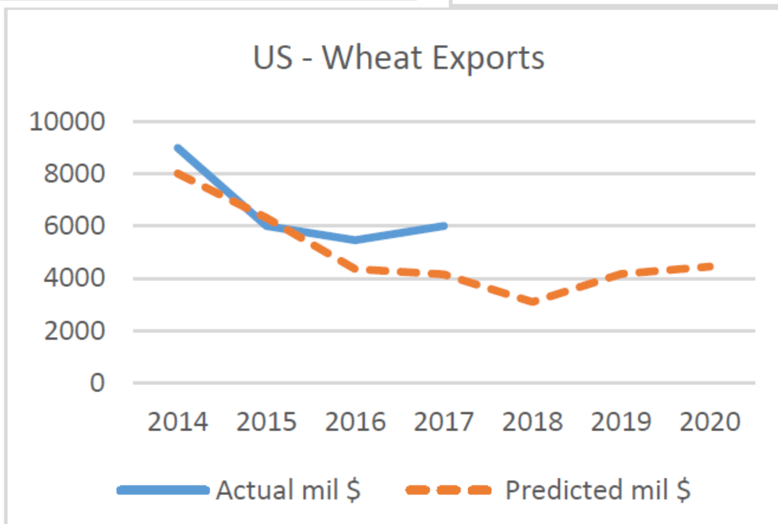
Unsupervised Model Predictions







Using DL



- Neural Networks
- Boosting
- Bagging
- ARIMA
- Linear Regression

So what?

- Existing forecasts: by WTO, OECD and USDA - ***model-based analyses*** and ***expert judgement*** with ***high variability***
  - Farmers likely consider the potential demand from alternative foreign sources before deciding to plant crops, especially in large exporters!
  - Countries setting budgets for farm programs need better predictions of prices and trade flows for assessing domestic production and consumption needs

- Offer an alternative to complex trade models and expert judgment analyses by relying on data-driven and deep learning approaches that allow for robust specifications of complex economic relationships

Ongoing work on Association Rules (substitutes and complements in international trade) and Ensemble Machine Learning (G20 versus WTO Policy making)