Article

International Trade Analysis using EGM Neural Networks and Random Forests

Abstract

Analysing the International Trade using Extended Gravity Model, Neural Networks and Random Forest. We set to Analyse tradeflow patterns using both Intuitive and Extended Gravity Model. Further Implemented other Gravity Model Estimators like Pseudo Maximum Likelihood Estimator (PPML) And GPML. Implementation of Neural Network and Random Forest on Data from 1948-2019 containing 18 Countries and Comparison of predictive powers using R^2. The Results Signify that Random Forest is better than Neural Network, Fixed Estimation, PPML and OLS. Analysing Trade Pattern between India and USA between 2014-2019 estimated using Random Forest ending the Study by suggesting the Future Research Areas,

This paper is prepared by:

Prakhar Pradhan

Aditya Gupta

Raj Aryan

Saral Verma

Atul Umak

Chirag Sharma

Kaushal Chaudhary

Sumit Singh

Sandeep Gautam

Anurag Yadav

Administrative Support:

Professor Somesh Kumar Mathur

April 2021

The Authors Want to Thank **Professor Somesh Kumar Mathur** for his Guidance throughout the journey.

INTRODUCTION

The Gravity Model is the workhorse of International Trade Literature. It all started with Tinbergen (1962), relating trade flows with Economic sizes and Geographical Distance between two Trading partners. It is a crucial tool to find Trade-Related effects. Since then, there are many versions of the Gravity model, From Traditional Gravity Theory to Extended Gravity Theory. In the past, The Gravity Model used to be applied to Trade in Goods only. Still, recently it is applied to Trade in Services (Kimura 2006).

Nowadays, Structural Gravity Models (i.e., Theoretically Grounded) are coming into existence. This paper will analyze both the intuitive and extended gravity models in forthcoming sections as the intuitive model does not account for multilateral resistance terms. To estimate the comprehensive gravity model incorporating the multilateral resistance terms, we move on to Fixed Effects Estimation.

Further to deal with "Heteroskedasticity" and zero trade flow values, we will apply Poisson Pseudo-Maximum Likelihood Estimator because the Intuitive Model uses OLS, which drops zero Trade values and may lead to inappropriate results. Moreover, we will also use an extension of PPML, GPML developed by Anderson (2015), to analyze the effects of different trade Flow parameters.

Moving on, we will deploy Regression Algorithms to analyse the Trade data. We will use Neural Networks and Random Forest Regression to analyse the panel data. We will compare the model's predictive powers by famous evaluation metric R2 square. Sometimes there is a tendency for the model to get over-fit, for which We will deal with regularisation and hyper-tuning.

Moreover, there are multiple tools available for increasing prediction accuracies like Bagging, Random Forest, Boosting, and Extra-trees Regression. We will go with Random Forest Regression to improve the prediction accuracy and to deal with overfitting. We will analyse the trade panel data from 1948-2019 by various econometric models and Regression Algorithms and compare their predictive powers in the forthcoming sections.

Literature review

Tinbergen (1962) gave the Gravity Model followed by Leamer and Levinsohn's (1995) 's statement that the gravity model gave many robust findings in empirical economics. Anderson and Wincoop (2003) gave the "gravity with gravitas "model, a demand function that incorporates both outward and inward multilateral resistance terms. An intuitive model does not account for these terms. Trade cost in the intuitive model is a function of distance only. In the Structural Gravity model, we define Trade cost as a function of distance, common Border, colony, common language, and many agreements like EU, WTO, and RTA. To deal with Multilateral resistance terms, we can deploy two ways: fixed effect estimation and the approximation technique developed by Baier and Bergstrand (2009).

Silva and Tenreyro (2006) developed a way to deal with Zero Trade Flow values and Heteroskedasticity as it runs like nonlinear squares on the equation. Researchers believed it is better than OLS as it performs better even in the dataset with many zeros, and OLS drops all those values. Anderson (2015) developed the extension of PPML, GPML and many have started using Poisson because it outperforms the OLS. Head and Mayer (2014) highlighted the importance of the fixed effect.

By using country-specific selected products, all monadic terms, including multilateral terms, are captured. Wohl and Kennedy (2018) examined that neural networks have the nonlinear advantage to analyze complex relationships within the dataset. They Pointed that NN is more robust and highly adaptive. Baxter and Hersh (2017) and Storm, Baylis (2019) offered the study of Time Series projections by applying a deep learning approach as it was being used to study trade disruptions in the world like Brexit and the US-China. Ke et al.2017 showed R-square as the Statistical Measure among both Supervised and Unsupervised Learnings.

Dataset preview

This Study collects panel dataset from the centre d'Etudes et d'Informations Internationales (CEPII), UN-Comtrade database. It includes Trade Flow and GDP between countries which are in real terms, not in nominal terms, as Multilateral words affect these. Other variables to account for Trade Cost like Geographical distance, Common Language, religion, Border. Many agreement variables like RTA, GATT, EU, WTO are also accounted for in this study. Our study Consists of Top 18 Countries: India, China, USA, Japan, Singapore, Belgium, Hongkong, Bangladesh, Pakistan, France, Italy, Korea, Saudi Arabia, United Arab, Germany, Myanmar, United Kingdom, and Switzerland. We will analyse trade data from 1948-2019, so a total of 23328 rows and 25 columns are present in our study. The different Section involves implementing the Intuitive and Extended Gravity model followed by PPML, GPML. Finally, we will analyse trade data using Neural Network and Random Forest and compare these models' predictive powers using R2 Square.

Data Cleaning

Data Cleaning is an important task to avoid deviations in the data and prevent wrong predictions. Firstly, we have removed all Zero Trade Flow Values as it may cause inconvenience, and secondly, we have extrapolated the missing data using forward and backward linear extrapolation. Then we have standardized the data by scaling them so that the mean is zero and the standard deviation is one.

$$X_s = \frac{x_i - mean(\varkappa)}{stdev(x)}$$

Where μ is the mean defined, and σ is the Standard Deviation.

$$\mu = \frac{1}{N}\sum_{i=1}^{N} x_i$$

$$\sigma = \frac{1}{N}\sum_{i=1}^{N}(x_i - \mu)$$

N is the number of observations observed, and X_l is a Random Variable.

Table 1: Description of Variables.

Variables	Description
distwces	Using CES formulating Population-weighted distance between populated cities.
Gdp_o / gdp_d	GDP (nominal) of the origin or destination country (current thousands US$).
Iso3	Country alphabetical and Numerical Code
Contig	Equal to 1 if countries shared common Border.
comcol	Dummy Column equal to 1 if countries share a joint colonizer post-1945.
comrelig	Religious proximity index
Rta	Equal to 1 if the country has an RTA agreement.
Rta_type	Indicates the type of RTA (customs union, for instance)
Rta_coverage	Equal to 1 if RTA covers both Goods and Services.
Gatt_O / Gatt_d	Dummy equal to1 if origin or destination country currently is a GATT member.
Wto_o / Wto_d	Dummy equal to 1 if origin or destination country currently is a WTO member.
Eu_o / Eu_d	Dummy equal to 1 if origin or destination country currently is an EU member.
Comlang_off	Dummy equal to 1 if countries share a familiar official or primary language.

Comlang_ethno	Equal to 1 if at least 9% of the population share the common language.
entry_cost_o / entry_cost_d	Cost of business (% of GNI per capita) of origin or destination country.
Tradeflow_comtrade_o	Trade flow as reported by the exporter (in thousands current US$) (source: Comtrade)

Model method

Gravity Model:

The Basic Gravity Model or the Traditional Gravity Model is as follows:

$$\log X_{ij} = c + b_1 \log(GDP_i) + b_2 \log(GDP_j) + b_3 \log(T_{ij}) + e_{ij}$$

$$\log(T_{ij}) = \log(distance_{ij})$$

where X indicates exports from the country I to j. GDP is each country's Gross Domestic Product, and T_{ij} represents Trade costs between the countries which is a function of distance. E_{ij} refers to the error term.

Fixed Effect Estimation is done by modifying Anderson and Wincoop's (2003) equation.

$$Log(X_{ij}) = -\log(Y) + \log(Y_i) - \log(\pi_i) + \log(Y_j) - \log(P_j) + (1 - \sigma)(\log(T_{ij}))$$

Log(Y) is the constant term equal to GDP theoretically, but it can be considered consistent across all exporters and importers for estimation purposes.

$$Log(T_{ij}) = b_1 \log(distwces_{ij}) + b_2 contig_{ij} + b_3 comlang\_off_{ij} + b_4 comlang\_ethno_{ij} + b_5 rta_{ij} + b_6 rta\_coverage_{ij} + b_7 rta\_type_{ij} + b_8 gatt\_o_i + b_9 gatt\_d_j + b_{10} eu\_o_i + b_{11} eu\_d_j + b_{12} wto\_o_i + b_{13} wto\_d_j + b_{14} comcol_{ij} + b_{15} comrelig_{ij} + b_{16} entry\_cost\_o_i + b_{17} entry\_cost\_d_j$$

PPML equation is written by considering the nonlinear form of Anderson and Wincoop's equation accounting for the multiplicative error term.

$$Log(X_{ij}) = \log(Y_I^K) + \log(E_j^k) - \log(Y^K) + (1 - \sigma_K)[\log(T_{ij}^k) - \log(\pi_i^k) - \log(P_j^k)] + \log(e_{ij}^k)$$

Neural Network

Neural Network is a computational deep learning model that works similar to a Biological neural network in the human brain. It is successfully applied to a variety of problems ranging from pattern recognition, natural language processing, etc. it is an expert in analyzing the complex patterns in the dataset by learning in such a way that it compares its predicted outputs with the real ones and adjust its weight by minimizing the error. In backpropagation networks, the network starts with random weights and compares the outcomes with actual ones, and weights are changed according to the error. The cycle is repeated until the average error converges to an optimum. The equation is written as follows:

$$Y = f(w_1x_1 + w_2x_2 + b)$$

where w_i represents the weights and X_I represents the inputs. In our research X accounts for various parameters like distance, contig, comlang_off, comlang_ethno, rta, rta_type, rta_coverage, gatt, eu,wto, comcol and comrelig.

Different Activation are there like tanh, sigmoid and Relu. Sigmoid Function is defined as follows:

$$\text{sigmoid}(x) = 1/(1 + \exp(-x))$$

Figure 1 represents the Forward Neural Network. X_i represents the inputs and $\hat{y}$ The output. In our study, inputs are various parameters given in Table 1.

Random Forest

Random Forest Regression is written in the following form: Supervised learning model, which prevents overfitting the data. It uses multiple decision Trees, which increases accuracy. It works in a way by creating numerous decision Trees by using a Technique called 'Bagging.' Figure 2 shows the typical depiction of Random Forest. Output is obtained by taking the mean of all predictions from each tree. Hyper Tuning Parameters in Random Forest to improve accuracy are Max depth and n_estimators.

Typical equation of Random Forest is as follows:

$$f_{rf}^{B}(x) = (1/B)*(\sum^{B}_{b=1} T_b(x))$$

$$\text{Var}((1/B)*(\sum^{B}_{b=1} T_b(x))) = \rho\sigma^2 + \sigma^2((1 - \rho)/B)$$

Here B is the number of trees T_B = Random Forest Tree. ρ is the positive correlation. B can be calculated by cross-validation or by estimating out-of-bag error. Bootstrapping procedure is followed as it decreases the variance

As B increases, variance decreases by decreasing the correlation between the trees. Each growing tree on bootstrapped data before each split selects m<p at random of the input variables. After B trees $\{T(x;\ \theta_b)\}_1^B$ is increased, Random Forest Predictor equation can be written as follows:

$$\hat{f}^B(x) = \frac{1}{B}\sum_{b=1}^{B} T(x;\ \theta_b)$$

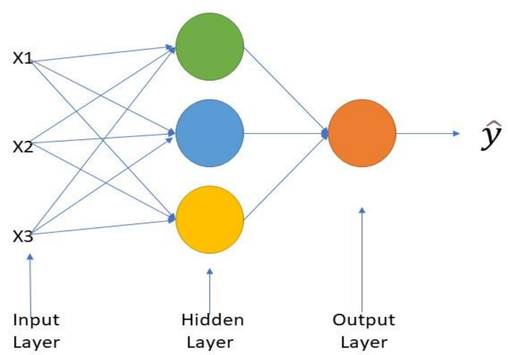


Figure 1: Neural Network consists of Input, hidden, and output layers. $\hat{y}$: Output.

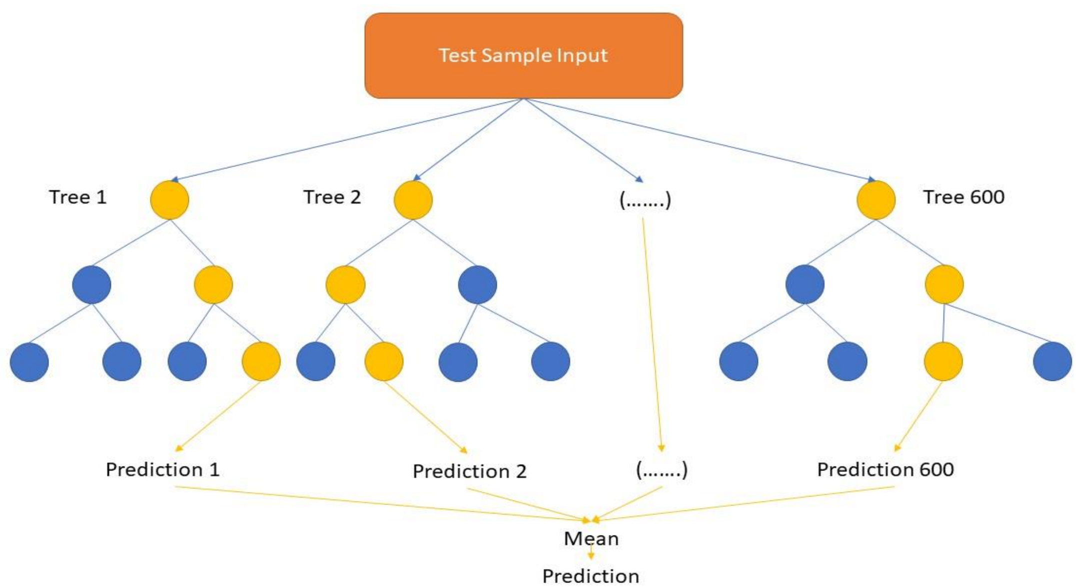


Figure 2: Random Forest Depiction. N is the number of trees, and output is predicted using the mean of all Output of Trees.

Evaluation and results

To evaluate our econometric models, we use Hypothesis testing. Generally, distance has a negative Coefficient with a tremendous negative t-score, and the null hypothesis that distance has no relation with the Trade flow is rejected. In OLS estimation, we use Robust Standard error, and distance is used as the cluster. R^2 Square value, which is Coefficient of Determination, and RMSE is Root mean square error. Generally, the value of R^2 Square is between 0 and 1. Value of 0 predicts that the model is failed, and a weight of 1 predicts model perfectly fits Data. RMSE is the square root of differences between predicted and actual values.

R^2 is the Coefficient of determination which is defined below:

$$R^2 = 1 - \frac{SS_{res}}{SS_{TOT}} = 1 - \frac{\sum_i (y_i - \hat{y})^2}{\sum_i (y_i - \bar{y})^2}$$

Where SS_{res} and SS_{tot} represent the residual sum of squares and total sum square. Y_i is the input vector, $\hat{y}$ is the output vector and $\bar{y}$ The mean of observed data.

Moreover, coefficients of different variables in OLS, PPML and others model gives their effect on trade. To check the statistical significance, we use the p-value as our measure. In Random Forest, several trees are being used to improve the accuracy boosting parameters like learning rate. In Neural Network Analysis, Feed Forward Algorithm, we have used Rprop Algorithm to minimize loss function and other hyper tuning parameters like learning rate, Number of epochs, and Number of hidden layers.

Table 2: Variables effect on Trade flow using Econometric Models.

Variable	OLS	Fixed Effects	PPML	GPML
log(distwces)	-0.548	-0.6	-0.582	-0.621
contig	-0.086	0.196	0.339	0.202
comcol	0.281	-0.002	0.47	-0.057
comrelig	0.074	0.853	0.443	0.75
rta	-0.809	-0.11	-0.012	0.041
rta_type	-0.001	-0.054	-0.126	-0.049
rta_coverage	0.429	0.13	0.139	0.082
gatt_o	0.39	0.636	0.854	0.961
gatt_d	-0.135	0.61	0.475	0.596

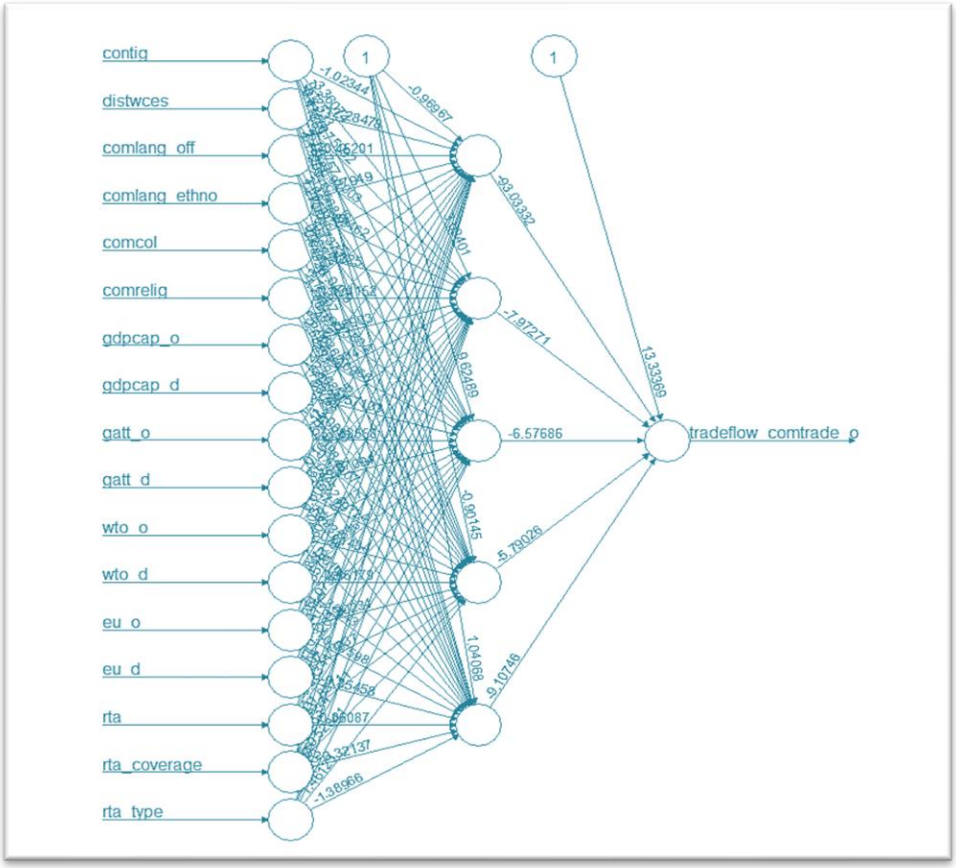
wto_o	-0.307	0.47	1.01	0.702
wto_d	-0.755	0.486	0.652	0.48
eu_o	0.022	1.132	0.901	0.597
eu_d	-0.063	-0.262	-0.265	-0.348
comlang_off	-0.005	-0.101	-0.113	-0.077
comlang_ethno	0.661	0.125	0.143	0.039
entry_cost_o	-0.019	-0.009	-0.001	-0.0061
entry_cost_d	-0.013	-0.005	-0.003	-0.004

We are observing table 2 that the coefficients of different Trade flow parameters using other econometric models. For say, distance and entry costs are negatively affecting the trade flow. Moreover, being an EU, WTO, or GATT affects trade positively. Now below table presents the prediction power of OLS, Fixed Effects, Neural Network, and Random Forest.

Table 3: Model vs Predictive Power (R^2)

Model	Prediction power (R^2)
OLS	0.71
Fixed Effects	0.87
Neural Network	0.85
Random Forest	0.96

From table 3, we can conclude the predictive powers of these models using a famous evaluation metric used for Regression, i.e., R^2 square. We can see that for the OLS model, it is 0.71, and for the fixed effect, it is 0.87. the increase in this accuracy accounts for incorporating the multilateral resistance terms. Further, for the neural network, it is 0.85, and for the random forest, it is 0.96. This clarifies that Machine Learning Algorithms help understand the complex relations within the dataset quickly and can be used to predict trade patterns as we can see random forest model predicted values closely followed the actual trade patterns.

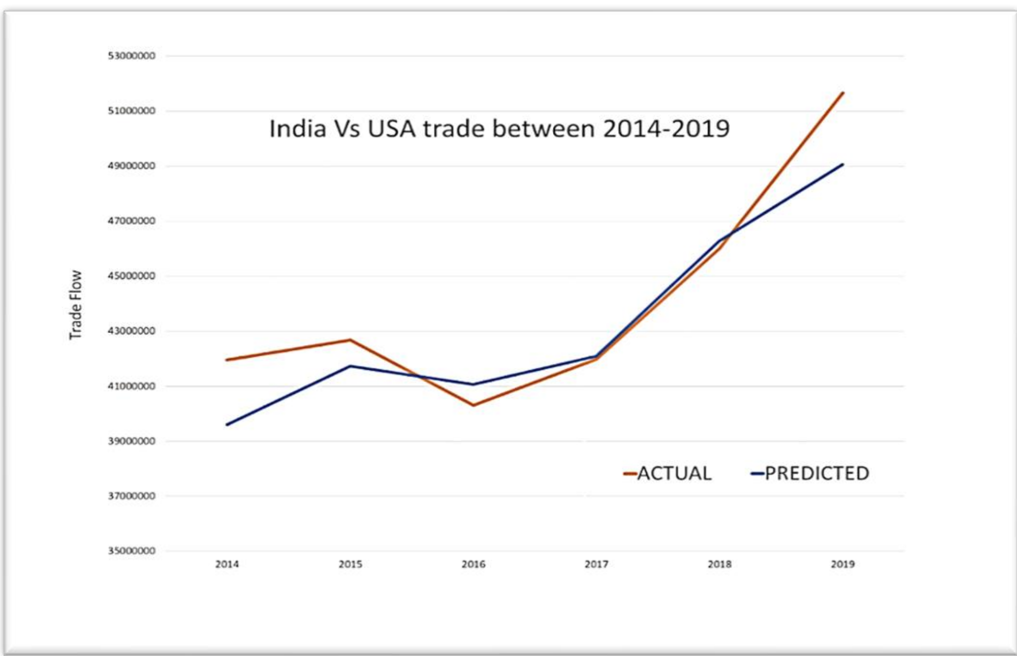


Hidden Layers: 5

Algo to minimise loss function: “Rprop”.

Dense Neural Network Depiction

Hyper Tuning Parameters: Number of Hidden layers



Random Forest Regressor Prediction. Trade of India VS USA between 2014-19

Conclusion

Neural networks and Random Forest have a high degree of accuracy in prediction than the R-Square value of the Traditional Gravity model. The model's explanatory power increased by using the Fixed effect model, and it underlines the importance of multilateral resistance terms. PPML is more preferred to Other econometric methods.

Random Forest RMSE is much lower than both Neural Network and gravity ones which indicates that in the future, we should move on using the machine learning models for estimating global trade patterns as it helps in dealing with large datasets which have numerous interconnected variables. We saw that the Intuitive Gravity Model is a basic one that has an R square value of 0.71. By applying the fixed effect, including multilateral terms, we got to see the R square value increased to 0.87.

We also used new methods like PPML and GPML and saw various parameters on Trade flow. Finally, we applied the neural network and Random Forest Regression, and we saw the R-square value to be more than an intuitive Gravity Model that clearly states the model is better. We noticed that Trade Pattern between the USA and India predicted values closely followed the actual values, signifying that We should move on to Supervised and Unsupervised Learning Methods to Analyse Trade Patterns. This will be a beneficial aid to policymakers and firms engaged in the international trade business and provide the impulsion to measure and compute the effects of Trade Collaboration.

Factors like Common Border, language, everyday ethics play a significant role in trade. Also, being a member of WTO or GATT, or EU plays a vital role in Trade flow. Further Estimation of Aggregate Trade data using a neural network model offers various ways to measure its impact, effects, and possible ways to sustain and maintain the very Trade Relations with any country. Both supervised and Unsupervised Algorithms provide valuable insights and a high degree of accuracy and may help in time of uncertainty and will help In a technical analysis as data have driven estimations has risen in recent years and will change the patterns of comparative advantage and will affect the length of Global Value Chain.

Artificial Intelligence will help reduce trade, transaction, and information costs and increase Trade flow in the future.

Future Research

Machine Learning methods demonstrated great Computational power and prediction accuracy in the econometric world. New areas of research can include the use of CGANs (Conditional GAN). A Generative Adversarial Network(GAN) uses Two Neural networks: The generator, and the second is Discriminator. Conditional GAN is the version of GAN which is used for Regression. There is a lot to explore in the future, incorporating machine learning and deep learning approaches.

References

Wohl, Isaac & Kennedy, Jim. (2018). Neural Network Analysis of International Trade.

Dumor, K.; Yao, L. Estimating China's Trade with Its Partner Countries within the Belt and Road Initiative Using Neural Network Analysis. Sustainability 2019, 11, 1449.

Bikker, Jacob. (2009). An extended gravity model with substitution applied to international trade. The Gravity Model in the International Trade: Advances and Applications.

Machine Learning in Gravity Models: An Application to Agricultural Trade, Gopinath, Munisamy and Batarseh, Feras A and Beckman, Jayson

Mullainathan, Sendhil, and Jann Spiess. 2017. "Machine Learning: An Applied Econometric Approach." Journal of Economic Perspectives, 31 (2): 87-106.

Yoto V. Yotov, A simple solution to the distance puzzle in international trade, Economics Letters,2012

Yotov, Yoto. "Gravity Training Course." USITC, August 2016.

Anderson, James and Eric van Wincoop. "Gravity with Gravitas: A Solution to the Border Puzzle." American Economic Review, vol. 93, no. 1, March 2003.

Santos Silva, J. M. C. and Silvana Tenreyro. "The Log of Gravity." Review of Economics and Statistics, vol. 88, no. 4, November 2006.

Tkacz, G.; Hu, S. Forecasting GDP Growth Using Artificial Neural Networks; Bank of Canada: Ottawa, ON, Canada, 1999.

Tillema, F.; Van Zuilekom, K.M.; Van Maarseveen, M.F. Comparison of neural networks and gravity models in trip distribution. Comput. Aided Civ. Infrastruct. Eng. 2006, 21, 104–119.

Anderson, J. 1979. "A Theoretical Foundation for the Gravity Model." American Economic Review, 69(1): 106-116.

Arvis, J.-F., and B. Shepherd. 2013. "The Poisson Quasi-Maximum Likelihood Estimator: A Solution to the Adding Up Problem in Gravity Models." Applied Economics Letters, 20(6): 515-519.

Eaton, J., and S. Kortum. 2002. "Technology, Geography, and Trade." Econometrica, 70(5): 1741-1779.

Kimura, F., and H.-H. Lee. 2006, "The Gravity Equation in the International Trade in Services."

Anderson, J., M. Larch, and Y. Yotov. 2015. “Estimating General Equilibrium Trade Policy Effects: GE PPML.” CESifo Working Paper No. 5592.

Tinbergen, J. 1962. Shaping the World Economy: Suggestions for an International Economic Policy. New York: The Twentieth Century Fund.

Leamer, E., and J. Levinsohn. 1995. “International Trade Theory: The Evidence” in G. Grossman and K. Rogoff (eds.) Handbook of International Economics. Amsterdam: Elsevier.

Head, K., and T. Mayer. 2014. “Gravity Equations: Workhorse, Toolkit, and Cookbook” in G. Gopinath, E. Helpman, and K. Rogoff (eds.) Handbook of International Economics Vol. IV, Amsterdam: Elsevier.

Baier, S., and J. Bergstrand. 2009 “Bonus Vetus OLS: A Simple Method for Approximating International Trade Cost Effects using the Gravity Equation.”

Ke, Guolin, Qi Meng, Thomas Finley, Taifeng Wang, Wei Chen, Weidong Ma, Qiwei Ye and Tie-Yan Liu (pp. 1-9), “LightGBM: A Highly Efficient Gradient Boosting” in (Long Beach, CA: 31st Conference on Neural Information Processing Systems, NIPS, 2017)

Baxter, Marianne, and Jonathan Hersh. Robust Determinants of Bilateral Trade. Paper Presented at Society for Economic Dynamics, 2017.

Storm, Hugo, Kathy Baylis, and Thomas Heckelei, “Machine Learning in Agricultural and Applied Economics,” European Review of Agricultural Economics, 2019.