# Analysis of Bilateral Trade Flow and Machine Learning Algorithms for GDP Forecasting

Jingwen Sun
School of International and Public Affairs, Columbia University,
New York, USA
jingwensun321@gmail.com

Yuan Suo
School of Social Sciences,
University of California,
Irvine, USA
suoyuandl@163.com

Seeha Park
Hankuk Academy of
Foreign Studies
Yongin, South Korea
seehalemon@gmail.com

Tianze Xu
College of Liberal Arts and Sciences,
University of Connecticut, Storrs, USA
constans_x@163.com

Yizhu Liu
Pius XI Catholic High School,
Milwaukee, USA
yizhuliupius@yeah.net

Weiqi Wang
Lawrence Woodmere Academy,
New York, USA
weiqiwanglwa@163.com

*Abstract*—**The terms imports and exports describe goods and services traded between countries. Countries import goods they cannot produce domestically or can obtain at a lower cost from another country. According to the World Trade Organization (WTO) reports, the U.S. is the world's largest importer based on capital investment, followed by the E.U., China, Germany, and Japan. For exports, China leads the world with an official trade amount of $1.904 trillion in 2013. E.U. ranks second, followed by U.S., Germany, and Japan. Trade in goods and services is defined as a change in ownership of material resources and services between economies. Trade indicators include the sale of goods and services as well as barter transactions or goods exchanged and are measured in million USD, the percentage of GDP for net trade, and the annual export and import growth. This study analyzes imports and exports of all countries for the 1960-2017 period and evaluates the correlations in trade statistics to predict future imports and exports. Since the GDP for any country depends mainly on trade, this study examines trade data and compares various machine learning algorithms to forecast a country's GDP.**

***Keywords-imports; exports; GDP; trade statistics; GDP forecast***

## I. INTRODUCTION

Trade data include exports, imports, and trade balance. In general, international trade represents the economic activity of a country related to some economic relationship with another country. A series of such activities forms a country's trade balance. The trade volume of a country indicates the country's collective effect of macroeconomic policies. Analyzing this effect from international trade policies requires data on the country's exports and imports and the long-run equilibrium relationship between these two variables. Here, there always is the effect of time lag on trade volume since any change in a country's import/export demand does not happen quickly, and therefore, trade data require a deep analysis [1]. One factor in explaining trade remains the comparative advantage. However, some new factors such as consumer preferences, advantages in the economy of scale, and the use of global production have emerged, and such factors may alter international trade patterns. In addition, bilateral trade is important in understanding the consequences of international trade [2].

Economies that are oriented mainly toward trade must carefully analyze their bilateral trade data since international trade flow and its direction, composition, and linkages require bilateral as well as multilateral analyses. Therefore, data should always be carefully examined in any empirical analysis [3]. Several studies have analyzed exports, imports, the economy, and their relationships, including cointegration between exports and imports [4, 5], long-run trade elasticity in less developed countries [6], cointegration in specific groups of countries [7], and related technical issues [8]. Some studies have investigated the importance of exports and imports in forming the economy and improving the quality of life [9]. There are various methods for determining the importance of exports and imports in the economy and analyzing their impacts on economic growth. One uses simple (and sometimes multiple) regression between these three variables and other factors, and another employs causality, determining the factors and causes as well as inverse causality. Some studies have used VAR and VEC models to exclude this causality problem. A cointegration test has verified a long-run equilibrium relationship between GDP, exports, and imports in Tunisia [10], finding these two variables to influence economic growth and thus highlighting their importance. A study of the Nigerian economy has suggested economic growth to be determined by exports, imports, labor, and the exchange rate and posited their cointegration, also showing positive relationships of exports and labor to economic growth and negative effects of imports and the exchange rate on the economy [11]. A study of the Ethiopian economy has revealed that the growth rate of real exports has a positive relationship with the rate of economic growth, analyzing that, although the effect is non-significant in the short run, there is a strong long-run relationship between these two variables [12]. These studies show that exports and imports have great impacts on economic growth. Economic growth is typically represented with GDP, GDP per capita, trade balance, consumption, investment, and government spending.

In addition to the impact on economic growth, exports and imports are important for other reasons. An example is the model of innovation and trade, which predicts positive shocks in exports can facilitate more productive and innovative firms. Here the accompanying rent from the innovation effort of companies increases with the firm's market size, known as the market size effect [13]. A country's GDP is a primary indicator of the country's economy. GDP represents the worth (in dollars) of all goods and services produced and imported/exported over a period of time. This is measured relative to previous GDP. An increase in GDP shows the growth of a country's economy, and this figure is considered a key benchmark for the country's economy. The U.S. has the largest GDP in the world, followed by China, Japan, Germany, the U.K., France, and India. In most cases GDP measurement is complicated. In simple terms, GDP is the income earned (gross profit) and expenditure spent by a country, and since it is considered relative to countries, exports and imports play a vital role in deciding a country's GDP. GDP can be calculated as the total spending on all final goods and services (Consumption of goods and services (C) + Gross Investments (I) + Government Purchases (G) + (Exports (X) - Imports (M)), i.e. GDP=C+I+G+(X-M).

## II. World Trade Data Statistics

Trade statistics represent a unique dataset for modern economies. Trade data statistics can indicate economic geography and provide insights into economic development and globalization. International trade plays an important role in a wage-based economy. International trade data provide insights into global relationships between countries as well as information on the relationship between local and regional economies [14]. Statistical data confirm that trade has continued to support economic growth and development, helping to reduce poverty around the world. World merchandise exports have increased in value by about 32% since 2006, reaching USD 16 trillion in 2016. At the same time, world exports of commercial services have accelerated by about 64%, reaching a total of USD 4.77 trillion [15]. The importance of world trade data statistics is clear, but its usefulness depends on various factors such as data availability, analysis requirements, and data accuracy. Therefore, measuring trade statistics must adhere to the Standard International Trade Classification [16]. This is important since insufficient data or analysis may cause misinterpretation of variables. According to the World Fact book, the world's top exporter is China with the total export of $2.157 trillion in 2017, followed by the E.U. with $1.929 trillion in 2016 and the U.S. with $1.576 trillion in 2017 [17]. For top import countries, the World Fact book ranks U.S. as the leader with $2.352 trillion of total imports in 2017, followed by E.U. with $1.895 trillion in 2016 and China with $1.731 trillion in 2017 [18]. Germany and Japan followed. The total import for Germany was $1.104 trillion in 2017, and Japan was $625.7 billion in 2017. For exports, Germany exported $1.401 trillion in 2017, and Japan, $683.3 in 2017 [17, 18]. The top 18 world imports and exports can be seen in [19]. Cars top the list, with Germany as the largest exporter and the U.S. as the top importer. Refined petroleum follows, with the U.S as a sole leader in both exports and imports. The list also includes goods such as pharmaceuticals, gold, crude petroleum, telephone, broadcasting equipment, diamonds, petroleum gas, and aircrafts [20].

## III. Literature Review

Authors in [23] proposed a machine learning model for predicting agricultural commodity prices over one-, two-, and three-month periods ahead. They used the multivariate relevance vector machine based on Bayesian learning for regression and compared the performance of the MVRVM model to that of multiple-output artificial neural networks. Authors in [24] applied data mining to detect relationship patterns in customs administration data with market prices and current exchange rates in Ethiopia and discovered association rules to generate note worthy import/export patterns. They used datasets from the Ethiopian Revenue and Customs Authority, Central Statistics Agency, and the National Bank of Ethiopia and applied the WEKA tool for data analysis purposes, and the results verified that imported textile was significantly related to the market price and the currency exchange rate. They also concluded the Apriori algorithm as the fastest one in discovering association rules. Authors in [25] predicted bilateral trade flow, an important economic indicator, by using the gravity model of trade with a fully connected feed-forward neural network. They experimented with machine learning models by varying hidden layers and neurons in each hidden layer and found that fully connected feed-forward neural networks can improve the gravity model's prediction performance. They also proposed that the LSTM model may yield better results than fully connected neural networks for time series data. Author in [26] used machine learning and data mining techniques for publicly available commodity data and forecasted country GDP, finding a correlation between export-import data and GDP. He considered commodity trade and GDP as inputs to the algorithm and designed a model to predict GDP for another day with the given commodity trade for the new day. He used a multi-class support vector machine with a genetic algorithm based on a fuzzy set and artificial neural network to predict GDP. Authors in [27] implemented a back-propagation neural network based on a genetic algorithm for port throughput forecasting. They used a 12-year dataset such that 11 years of data was used for simulations/training and the final year was used for forecasting. They verified the proposed hybrid model, the GA-BP forecasting model, to show better accuracy but it took longer to converge.

Authors in [28] proposed a new machine learning approach for price modeling using a neural network with an advanced signal-processing tool. They used the proposed model to forecast prices of commodities such as coal, crude oil, and electricity and employed a mixture of a Gaussian neural network, showing significant improvements relative to other popular models. Authors in [29] constructed a GARCH model using an artificial neural network and evaluated its ability to forecast stock market volatility. They compared the performance of their model to that of other popular volatility models for various international stock indices. Authors in [30] proposed a model to predict future gold rates based on 22 market variables using machine learning techniques. They collected data from various online sources and implemented a linear regression model using artificial neural networks with the rapid miner tool. Authors in [31] proposed a unified modeling framework to justify the empirical regularity in the international trade network and analyzed the international trade network each year with exports of the country with other countries. They constructed a basic model with a directed weighted network for unified modeling.

## IV. MACHINE LEARNING IN GLOBAL TRADE DATA

Machine learning applies artificial intelligence (AI) to automatically learn and improve from experience without being specifically programmed. Machine learning focuses on the development of computer programs for accessing and learning from data. The process of learning begins with observations or data, including examples, direct experience, and instructions, to identify certain patterns in data for making better decisions. The main aim is to allow the computer to learn automatically without human intervention and engage in appropriate actions. AI and machine learning represent a good way for financial institutions to optimize margin valuation adjustment (MVA), which can be performed through the assistance of machine learning by reducing margins for derivatives through a combination of "executing pairs of offsetting derivative trades" and "executing offsetting strategies with the same dealer" [21]. By choosing the best combination of the initial margin, machine learning reduces trades in a given period of time, and the basis for this is the degree of initial margin reduction in the past from various combinations of trades [20]. Financial institutions emphasize cost-effective means for regulatory requirements, such as efficient trade execution, data reporting, and prudent regulation. In this regard, AI can be used to obtain information and process orders, and machine learning can help create "trading robots" that can respond quickly to market changes. Therefore, such innovations can be used by firms to estimate financial impacts more accurately and minimize trading costs [20].

## V. METHODOLOGY

### *A. Datasets*

The data for this research study is gathered from [32]. The data consisted of yearly import/export data from 217 countries for the 1960-2017 period. Data were collected from the BoP (Balance of Payments) statistics yearbook and presented in USD. Several missing data points were identified as a result of no substantial import/export participation by some countries during early years. Therefore, the data for the last 43 years were considered in this study (i.e., 1975-2017). The data containing the top 10 countries in imports and exports were extracted from the primary data source for the analysis. The data (in billion USD) were used for further analysis. Tables I and II present the import and export data, respectively, for the study period. Datasets show the U.S. as leading both imports and exports from 1975 to 2017. Therefore, it is evident that the U.S. had the highest GDP and the GDP was growing at a constant pace. For better understanding, the import and export datasets are plotted in Figures 1 and 2 respectively.

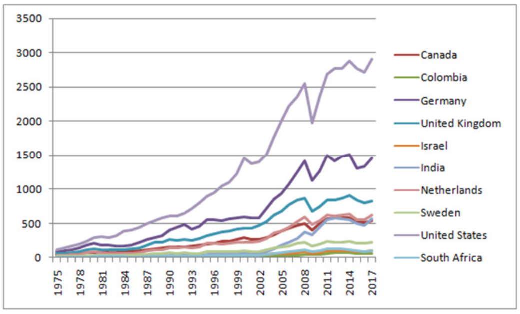


Fig. 1. Increase in imports.

TABLE I. IMPORTS (IN BILLON $) FOR 1975-2017

| Year | Canada | Colombia | Germany | UK | Israel | India | Netherlands | Sweden | US | S. Africa |
|---|---|---|---|---|---|---|---|---|---|---|
| 1975 | 42 | 3 | 94 | 64 | 7 | 7 | 40 | 21 | 120 | 12 |
| 1976 | 47 | 3 | 108 | 66 | 7 | 6 | 44 | 23 | 149 | 11 |
| 1977 | 49 | 3 | 123 | 74 | 8 | 7 | 51 | 24 | 180 | 10 |
| 1978 | 53 | 4 | 149 | 88 | 9 | 9 | 60 | 25 | 208 | 12 |
| 1979 | 63 | 4 | 193 | 115 | 11 | 12 | 76 | 35 | 248 | 15 |
| 1980 | 71 | 6 | 222 | 134 | 12 | 17 | 88 | 40 | 291 | 23 |
| 1981 | 79 | 7 | 195 | 122 | 13 | 18 | 76 | 36 | 310 | 26 |
| 1982 | 67 | 7 | 186 | 119 | 12 | 18 | 72 | 34 | 299 | 21 |
| 1983 | 74 | 6 | 181 | 118 | 12 | 18 | 70 | 32 | 323 | 18 |
| 1984 | 88 | 6 | 178 | 123 | 12 | 18 | 70 | 33 | 399 | 19 |
| 1985 | 93 | 6 | 184 | 128 | 12 | 19 | 73 | 35 | 410 | 13 |
| 1986 | 99 | 6 | 226 | 148 | 13 | 20 | 86 | 41 | 449 | 15 |
| 1987 | 107 | 6 | 271 | 183 | 17 | 23 | 104 | 51 | 501 | 18 |
| 1988 | 128 | 7 | 298 | 222 | 18 | 26 | 114 | 58 | 546 | 21 |
| 1989 | 141 | 7 | 319 | 233 | 18 | 29 | 119 | 62 | 580 | 21 |
| 1990 | 148 | 7 | 412 | 264 | 21 | 30 | 142 | 71 | 616 | 22 |
| 1991 | 152 | 7 | 452 | 251 | 23 | 28 | 146 | 66 | 610 | 22 |
| 1992 | 157 | 9 | 485 | 267 | 24 | 30 | 157 | 68 | 656 | 23 |
| 1993 | 168 | 12 | 420 | 255 | 27 | 31 | 144 | 56 | 713 | 24 |
| 1994 | 183 | 14 | 464 | 284 | 31 | 38 | 159 | 66 | 802 | 27 |
| 1995 | 199 | 16 | 559 | 327 | 36 | 49 | 210 | 81 | 891 | 34 |
| 1996 | 209 | 17 | 553 | 355 | 38 | 55 | 211 | 85 | 956 | 34 |

| Year | Canada | Colombia | Germany | UK | Israel | India | Netherlands | Sweden | US | S. Africa |
|---|---|---|---|---|---|---|---|---|---|---|
| 1997 | 237 | 19 | 537 | 380 | 38 | 59 | 206 | 85 | 1040 | 35 |
| 1998 | 241 | 18 | 565 | 396 | 36 | 60 | 217 | 90 | 1100 | 33 |
| 1999 | 259 | 14 | 579 | 419 | 41 | 63 | 223 | 91 | 1230 | 31 |
| 2000 | 287 | 15 | 595 | 440 | 47 | 74 | 231 | 97 | 1450 | 34 |
| 2001 | 268 | 17 | 587 | 439 | 44 | 72 | 231 | 87 | 1370 | 32 |
| 2002 | 271 | 16 | 589 | 471 | 43 | 76 | 243 | 91 | 1400 | 33 |
| 2003 | 295 | 17 | 726 | 529 | 45 | 93 | 286 | 113 | 1510 | 44 |
| 2004 | 337 | 21 | 858 | 625 | 53 | 131 | 363 | 134 | 1770 | 59 |
| 2005 | 385 | 26 | 934 | 686 | 59 | 182 | 392 | 149 | 2000 | 69 |
| 2006 | 430 | 31 | 1080 | 784 | 63 | 225 | 441 | 169 | 2220 | 84 |
| 2007 | 471 | 38 | 1250 | 841 | 75 | 279 | 522 | 204 | 2360 | 98 |
| 2008 | 508 | 46 | 1410 | 867 | 85 | 379 | 595 | 226 | 2550 | 108 |
| 2009 | 412 | 40 | 1130 | 677 | 64 | 328 | 486 | 167 | 1970 | 83 |
| 2010 | 500 | 49 | 1270 | 753 | 77 | 439 | 532 | 197 | 2350 | 103 |
| 2011 | 568 | 64 | 1500 | 840 | 93 | 553 | 615 | 233 | 2680 | 123 |
| 2012 | 587 | 70 | 1410 | 844 | 93 | 580 | 600 | 221 | 2760 | 124 |
| 2013 | 586 | 71 | 1480 | 869 | 92 | 560 | 618 | 224 | 2760 | 122 |
| 2014 | 585 | 76 | 1510 | 915 | 95 | 554 | 632 | 231 | 2870 | 116 |
| 2015 | 531 | 65 | 1310 | 840 | 85 | 492 | 552 | 200 | 2760 | 100 |
| 2016 | 513 | 55 | 1330 | 804 | 90 | 472 | 555 | 202 | 2710 | 90 |
| 2017 | 548 | 57 | 1460 | 838 | 97 | 561 | 619 | 222 | 2900 | 100 |

TABLE II. EXPORTS (IN BILLON $) FOR 1975-2017

| Year | Canada | Colombia | Germany | UK | Israel | India | Netherlands | Sweden | US | S. Africa |
|---|---|---|---|---|---|---|---|---|---|---|
| 1975 | 39 | 3 | 103 | 60 | 4 | 6 | 43 | 21 | 130 | 11 |
| 1976 | 45 | 3 | 119 | 63 | 5 | 7 | 47 | 22 | 142 | 10 |
| 1977 | 48 | 4 | 137 | 76 | 6 | 8 | 52 | 23 | 152 | 12 |
| 1978 | 54 | 4 | 166 | 92 | 7 | 9 | 60 | 27 | 178 | 15 |
| 1979 | 64 | 5 | 198 | 117 | 8 | 10 | 75 | 34 | 223 | 20 |
| 1980 | 75 | 6 | 217 | 146 | 9 | 12 | 87 | 39 | 272 | 29 |
| 1981 | 81 | 5 | 200 | 136 | 9 | 12 | 81 | 36 | 294 | 24 |
| 1982 | 80 | 5 | 201 | 128 | 9 | 13 | 78 | 34 | 275 | 21 |
| 1983 | 86 | 4 | 192 | 121 | 9 | 14 | 75 | 34 | 266 | 21 |
| 1984 | 100 | 6 | 193 | 122 | 10 | 14 | 76 | 36 | 291 | 20 |
| 1985 | 101 | 5 | 207 | 132 | 10 | 13 | 79 | 37 | 289 | 19 |
| 1986 | 102 | 7 | 272 | 144 | 12 | 14 | 92 | 44 | 310 | 20 |
| 1987 | 112 | 7 | 326 | 175 | 14 | 16 | 109 | 53 | 349 | 26 |
| 1988 | 132 | 7 | 358 | 191 | 15 | 18 | 122 | 60 | 431 | 27 |
| 1989 | 142 | 8 | 379 | 199 | 16 | 21 | 128 | 63 | 487 | 26 |
| 1990 | 149 | 9 | 457 | 239 | 18 | 23 | 153 | 71 | 535 | 28 |
| 1991 | 149 | 10 | 443 | 240 | 17 | 24 | 157 | 70 | 578 | 27 |
| 1992 | 155 | 10 | 473 | 254 | 20 | 25 | 169 | 72 | 617 | 28 |
| 1993 | 168 | 10 | 421 | 246 | 21 | 28 | 160 | 62 | 643 | 30 |
| 1994 | 189 | 11 | 468 | 277 | 24 | 32 | 178 | 74 | 703 | 31 |
| 1995 | 218 | 13 | 571 | 322 | 28 | 39 | 235 | 96 | 794 | 35 |
| 1996 | 233 | 14 | 573 | 352 | 30 | 41 | 236 | 101 | 852 | 36 |
| 1997 | 249 | 15 | 563 | 383 | 32 | 45 | 231 | 101 | 934 | 37 |
| 1998 | 253 | 14 | 594 | 384 | 33 | 46 | 240 | 105 | 933 | 35 |
| 1999 | 283 | 14 | 594 | 393 | 38 | 52 | 242 | 106 | 970 | 34 |
| 2000 | 328 | 16 | 601 | 409 | 47 | 60 | 247 | 112 | 1080 | 37 |
| 2001 | 310 | 16 | 622 | 401 | 41 | 63 | 248 | 105 | 1010 | 36 |
| 2002 | 304 | 15 | 680 | 421 | 40 | 71 | 260 | 107 | 979 | 37 |
| 2003 | 329 | 16 | 819 | 480 | 44 | 85 | 322 | 131 | 1020 | 48 |
| 2004 | 382 | 20 | 1000 | 563 | 54 | 116 | 410 | 161 | 1160 | 59 |
| 2005 | 431 | 25 | 1080 | 622 | 59 | 155 | 446 | 174 | 1290 | 69 |
| 2006 | 465 | 29 | 1240 | 719 | 63 | 193 | 499 | 199 | 1460 | 80 |
| 2007 | 502 | 35 | 1480 | 765 | 73 | 240 | 593 | 235 | 1650 | 94 |
| 2008 | 536 | 44 | 1640 | 781 | 84 | 305 | 673 | 257 | 1840 | 103 |
| 2009 | 392 | 39 | 1300 | 625 | 70 | 261 | 550 | 188 | 1580 | 84 |
| 2010 | 469 | 46 | 1440 | 689 | 82 | 348 | 603 | 220 | 1850 | 108 |
| 2011 | 547 | 64 | 1690 | 799 | 95 | 446 | 692 | 262 | 2130 | 127 |
| 2012 | 551 | 69 | 1630 | 792 | 93 | 444 | 679 | 249 | 2220 | 118 |
| 2013 | 556 | 68 | 1710 | 813 | 98 | 468 | 711 | 253 | 2290 | 113 |
| 2014 | 568 | 65 | 1780 | 854 | 100 | 486 | 727 | 257 | 2380 | 110 |
| 2015 | 492 | 46 | 1580 | 790 | 94 | 429 | 632 | 225 | 2260 | 97 |
| 2016 | 476 | 42 | 1600 | 749 | 97 | 430 | 641 | 225 | 2210 | 92 |
| 2017 | 511 | 48 | 1740 | 802 | 102 | 488 | 714 | 240 | 2330 | 104 |

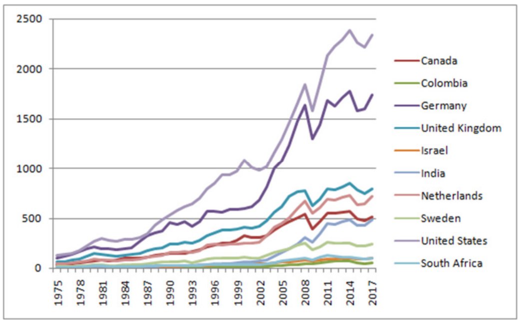


Fig. 2. Increase in exports.

With the growth in imports and exports of all countries identified, as shown in these plots, data analysis was conducted using linear regression to identify the correlation coefficient and other performance measures.

### B. *Data Analysis*

The GDP of any country is dependent on import and export data since these factors make the greatest contributions to the country's economy, and several algorithms have been implemented for the analysis of trade data. Therefore, this study focused on analyzing import and export data. The study initially employed linear regression, followed by a comparison with other machine learning algorithms. Here, preprocessed data from 10 countries were used as the input for linear regression, and the results for export and import datasets were obtained as shown in Tables III and IV, respectively.

TABLE III. EXPORTS ANALYSIS FOR TOP 10 COUNTRIES

| Country | CC | MAE | RMSE | RAE (%) | RRS(%) |
|---|---|---|---|---|---|
| Canada | 0.9642 | 40.606 | 46.7947 | 25.1067 | 25.8104 |
| Colombia | 0.8346 | 8.3626 | 10.8224 | 49.9367 | 53.1576 |
| Germany | 0.9375 | 162.9892 | 195.2004 | 32.1998 | 33.721 |
| India | 0.8353 | 78.3047 | 90.7998 | 54.2975 | 53.1019 |
| Israel | 0.9484 | 9.2247 | 10.3887 | 31.6044 | 30.7659 |
| Netherlands | 0.933 | 72.6982 | 84.1967 | 34.3541 | 34.8714 |
| S. Africa | 0.8842 | 13.7084 | 16.464 | 43.1783 | 45.1946 |
| Sweden | 0.942 | 22.7929 | 27.3736 | 31.1841 | 32.5885 |
| UK | 0.9639 | 61.4828 | 70.191 | 25.9338 | 25.899 |
| US | 0.9566 | 184.4345 | 211.1925 | 29.5354 | 28.3296 |

CC: Correlation Coefficient, MAE: Mean Absolute Error, RMSE: Root Mean Square Error, RAE: Relative Absolute Error, RRSE: Root Relative Squared Error

TABLE IV. IMPORTS ANALYSIS FOR TOP 10 COUNTRIES

| Country | CC | MAE | RMSE | RAE (%) | RRS(%) |
|---|---|---|---|---|---|
| Canada | 0.9585 | 44.3818 | 51.9594 | 27.3267 | 27.6743 |
| Colombia | 0.8578 | 9.0713 | 11.2877 | 48.4398 | 49.8184 |
| Germany | 0.9447 | 130.7667 | 154.7405 | 31.125 | 31.8091 |
| India | 0.8272 | 94.3616 | 110.4496 | 54.8779 | 54.241 |
| Israel | 0.9562 | 7.6478 | 8.748 | 29.0447 | 28.426 |
| Netherlands | 0.9347 | 62.5513 | 71.9923 | 34.2352 | 34.4388 |
| South Africa | 0.8703 | 15.8974 | 18.4622 | 46.8866 | 47.7568 |
| Sweden | 0.9344 | 20.9784 | 25.3766 | 32.8641 | 34.5429 |
| UK | 0.9628 | 67.9066 | 77.7684 | 25.9899 | 26.2868 |
| US | 0.9639 | 221.3444 | 249.555 | 25.9518 | 25.8817 |

CC: Correlation Coefficient, MAE: Mean Absolute Error, RMSE: Root Mean Square Error, RAE: Relative Absolute Error, RRSE: Root Relative Squared Error

From these import and export Tables, it can be observed that the correlation coefficient is close to 1, indicating a strong positive linear relationship. However, the RMSE is slightly high in some cases. Since the dataset consisted of only 43 instances (last 43 years of data), the linear regression model showed good correlations and high RMSE values. Therefore, further analysis using other machine learning algorithms was conducted with import and export datasets from the U.S. and Germany.

## VI. RESULTS AND DISCUSSION

To investigate on the best-performing machine learning algorithms for a better analysis of trade data and to forecast country GDP, this study considered 5 machine learning algorithms:

- Linear Regression (LR)
- RBF Regressor (RBF)
- Support Vector Machine (SVM)
- Regression by Discretization (RD)
- Reduced Error Pruning Tree(REP)

The U.S. and Germany import and export datasets were used as input datasets. A 10-fold cross-validation method was implemented to select training and testing datasets. Here correlation coefficients, root mean squared error, relative absolute error, and root relative squared error were considered as performance measures. A two-tailed test with a 0.05 confidence level was conducted and the results for these four performance measures are respectively shown in Tables V-VIII. Table V shows the correlation coefficients obtained for 4 datasets in the experiment with 5 algorithms. The Table shows that the RBF algorithm generated better models with high correlation coefficients, followed by RD, LR, SVM, and REP. Table VI (relative absolute error analysis) shows that RBF generated models with much smaller errors than LR and SVM. The results for RD lie between RBF and REP results. Therefore, RBF can be considered to perform well with respect to the relative absolute error measure.

TABLE V. CORRELATION COEFFICIENT ANALYSIS

| Data Set | LR | RBF | SVM | RD | REP |
|---|---|---|---|---|---|
| Germany_exports | 0.97 | 0.99 | 0.97 | 0.97 | 0.96 |
| Germany_imports | 0.97 | 0.99 | 0.97 | 0.97 | 0.96 |
| US_exports | 0.98 | 0.99 | 0.98 | 0.99 | 0.97 |
| US_imports | 0.98 | 0.99 | 0.98 | 0.99 | 0.97 |

TABLE VI. RELATIVE ABSOLUTE ERROR ANALYSIS

| Data Set | LR | RBF | SVM | RD | REP |
|---|---|---|---|---|---|
| Germany_exports | 35.46 | 13.78 | 35.85 | 20.99 | 22.45 |
| Germany_imports | 34.41 | 15.57 | 32.05 | 21.36 | 23.81 |
| US_exports | 34.21 | 12.02 | 33.16 | 16.60 | 22.30 |
| US_imports | 29.38 | 9.58 | 30.72 | 16.36 | 20.82 |

TABLE VII. ROOT MEAN SQUARED ERROR ANALYSIS

| Data Set | LR | RBF | SVM | RD | REP |
|---|---|---|---|---|---|
| Germany_exports | 187.53 | 86.55 | 195.54 | 118.76 | 139.53 |
| Germany_imports | 146.63 | 78.10 | 149.10 | 105.87 | 119.36 |
| US_exports | 205.34 | 83.24 | 237.78 | 111.97 | 155.62 |
| US_imports | 247.66 | 106.71 | 262.79 | 156.85 | 202.58 |

TABLE VIII. ROOT RELATIVE SQUARED ERROR ANALYSIS

| Data Set | LR | RBF | SVM | RD | REP |
|---|---|---|---|---|---|
| Germany_exports | 36.13 | 15.43 | 36.88 | 22.80 | 26.56 |
| Germany_imports | 33.85 | 16.69 | 33.01 | 23.69 | 26.95 |
| US_exports | 30.89 | 12.26 | 33.36 | 16.81 | 23.58 |
| US_imports | 28.26 | 11.45 | 29.47 | 17.66 | 23.25 |

Table VII (root mean squared error) and Table VIII (root relative squared error) also verify RBF as providing better results than the other algorithms. Although RMS values are high (with a smaller dataset), RBF obtains 60% better results than LR.

The performance plots for the above measures using the RBF algorithm are shown in Figures 3-6.

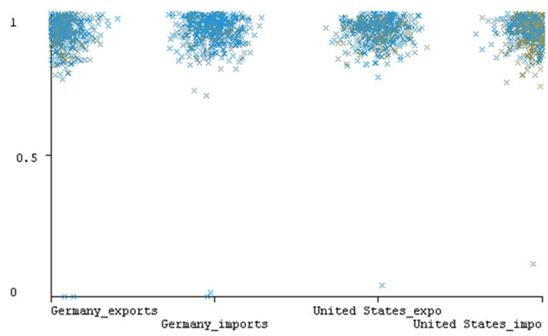


Fig. 3. Correlation coefficienet.

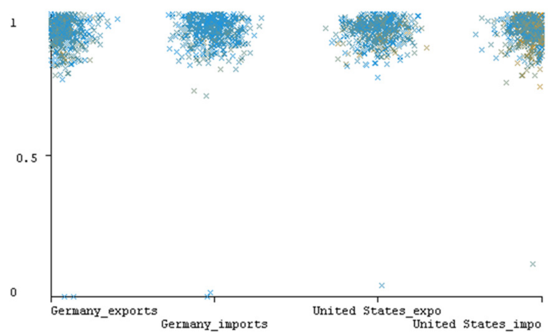


Fig. 4. Mean absolute error.

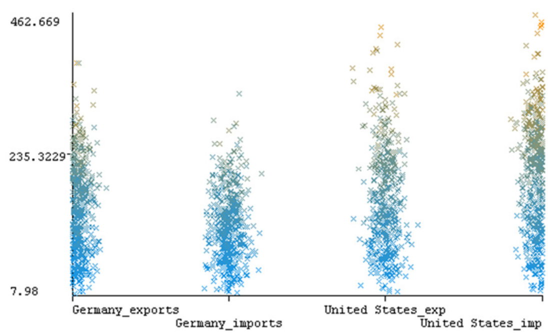


Fig. 5. Root mean squared error

Given these measures, the RBF algorithm is shown to provide good performance with a gain of about 60%. The results demonstrate that RBF outperformed the other four algorithms and that the models generated using RBF produced better results in forecasting of country GDP.


Fig. 6. Root relative squared error

## VII. CONCLUSIONS

Trade analysis results show a strong positive linear relationship in trade statistics. Imports and exports increased linearly with the world engaging in increased trade activity. In particular, the U.S., China, Japan, and India have made remarkable progress in the past decade in both exports and imports. This study takes two different directions. The first direction is the analysis of the top 10 import/export countries to show the correlation in trade data, and the second is to investigate different machine learning algorithms to identify the best algorithm for the prediction of trade data and country GDP. Here the WEKA data analysis tool was used for data analysis, and the experiments were conducted using five machine learning algorithms. The results show that all five algorithms exhibited good correlations but that the RBF algorithm outperformed the other algorithms. This suggests that the RBF algorithm may perform well in forecasting trade data to predict a country's GDP.

## REFERENCES

[1] J. Uddin, "Time Series Behavior of Imports and Exports of Bangladesh: Evidence from Cointegration Analysis and Error Correction Model", International Journal of Economics and Finance, Vol. 1, No. 2, pp. 156-162, 2009

[2] I. S. Kim, S. Liao, K. Imai, "Measuring Trade Profile with Granular Product-level Trade Data", available at: https://imai.fas.harvard.edu/research/files/BIGtrade.pdf, 2018

[3] R. Sen, "Analyzing International Trade Data in a Small Open Economy: The Case of Singapore", ASEAN Economic Bulletin, Vol. 17, No. 1, pp. 23-35, 2000

[4] M. Bahmani-Oskooee, H. J. Rhee, "Are imports and exports of Korea cointegrated?", International Economic Journal, Vol. 11, No. 1, pp. 109-114, 1997

[5] T. T. Cheong, "Are Malaysian exports and imports cointegrated? A comment", Sunway Academic Journal, Vol. 2, pp. 101-107, 2005

[6] M. Bahmani-Oskooee, "Cointegration approach to estimate the long-run trade elasticities in LDCs", International Economic Journal, Vol. 12, No. 3, pp. 89-96, 1998

[7] T. C. Tang, "Are imports and exports in the OIC member countries cointegrated? A reexamination", IIUM Journal of Economics and Management, Vol. 14, No. 1, pp. 1-31, 2006

[8] C. K. Choong, S. C. Soo, Z. Yusop, "Are Malaysian exports and imports cointegrated?", Sunway Academic Journal, Vol. 1, pp. 29-38, 2004

[9] M. J. Sirgy, D. J. Lee, C. Miller, J. E. Littlefield, E. G. Atay, "The Impact of Imports and Exports on A Country's Quality of Life", Social Indicators Research, Vol. 83, No. 2, pp. 245-281, 2007

[10] A .A. J. Saaed, M. A. Hussain, "Impact of Exports and Imports on Economic Growth: Evidence from Tunisia", Journal of Emerging Trends in Economics and Management Sciences, Vol. 6, No. 1, pp. 13-21, 2015

[11] D. Omotor, "The Role of Exports in the Economic Growth of Nigeria: The Bounds Test Analysis", International Journal of Economic Perspectives, Vol. 2, No. 3, pp. 222-235, 2008

[12] F. E. Chemeda, "The Role of Exports in Economic Growth with Reference to Ethiopian Country", Conference on Annual Meeting of American Agricultural Economics Association in Chicago, Chicago, USA, August 5-8, 2001

[13] P. Aghion, A. Bergeaud, M. Lequien, M. Melitz, The Impact of Exports on Innovation: Theory and Evidence, National Bureau of Economic Research, 2018

[14] L. Charles, G. Daudin, "Eighteenth-Century International Trade Statistics. Sources and Methods", Revue de l'OFCE, Vol. 4, pp. 7-36, 2015

[15] WTO, World Trade Statistical Review, WTO, 2017

[16] M. Jerven, "On the accuracy of trade and GDP statistics in Africa: Errors of commission and omission", Journal of African Trade, Vol. 1, No. 1, pp. 45-52, 2014

[17] The Central Intelligence Agency, The World Factbook, Country Comparison: Imports, available at: https://www.cia.gov/library/Publications/the-world-factbook/rankorder/2087rank.html

[18] The Central Intelligence Agency, The World Factbook, Country Comparison: Exports, available at: https://www.cia.gov/library/publications/the-world-factbook/rankorder/2078rank.html

[19] J. Desjardins, "The Top Importers and Exporters of the World's 18 Most Traded Goods", available at: http://www.visualcapitalist.com/top-importers-exporters-worlds-18-traded-goods/

[20] Teletrac Navman, "Top 18 Imports and Exports Around the World", available at: https://www.teletracnavman.com/infographics/top-imports-exports

[21] Financial Stability Board, Artificial Intelligence and Machine Learning in Financial Services: Market Developments and Financial Stability Implications, Financial Stability Board, 2017

[22] Info World, "An Intro to Genetic Algorithms", available at: https://www.infoworld.com/article/3151009/software/an-intro-to-genetic-algorithms.html

[23] A. M. Ticlavilca, D. M. Feuz, M. McKee, "Forecasting Agricultural Commodity Prices Using Multivariate Bayesian Machine Learning Regression", NCCC-134 Conference on Applied Commodity Price Analysis, Forecasting, and Market Risk Management. St. Louis, USA, April 19-20, 2010

[24] M. Manaye, B. Borena, "Association Pattern Discovery of Import Export Items in Ethiopia", HiLCoE Journal of Computer Science and Technology, Vol. 1, No. 2, pp. 82-88, 2013

[25] S. Circlaeys, C. Kanitkar, D. Kumazawa, "Bilateral Trade Flow Prediction", available at: http://cs229.stanford.edu/proj2017/final-reports/5240224.pdf, 2017

[26] H. R. Joseph, "GDP Forecasting through Data Mining of Seaport Export-Import Records", 9th International Conference on Data Mining, Las Vegas, USA, July 22-25, 2013

[27] F. F. Ping, F. X. Fei, "Multivariant forecasting mode of Guangdong province port throughput with genetic algorithms and Back Propagation neural network", Procedia – Social and Behavioral Sciences, Vol. 96, pp. 1165-1174, 2013

[28] M. Panella, F. Barcellona, R. L. D'Ecclesia, "Forecasting Energy Commodity Prices Using Neural Networks", Vol. 2012, Article ID 289810, 2012

[29] R. G. Donaldson, M. Kamstra, "An artificial neural network - GARCH model for international stock return volatility", Journal of Empirical Finance, Vol. 4, No. 1, pp. 17-46, 1997

[30] I. ur Sami, K. N. Junejo, "Predicting Future Gold Rates using Machine Learning Approach", International Journal of Advanced Computer Science and Applications, Vol. 8, No. 12, pp. 92-99, 2017

[31] S. Peluso, A. Mira, P. Muliere, A. Lomi, "International Trade: a Reinforced Url Network Model", available at: https://arxiv.org/abs/1601.03067, 2016

[32] World Bank, Indicators, available at: https://data.worldbank.org/indicator