

Predicting Level of Development for Different Countries

Zahoor Ahmad¹ & Aysha Saleem²

¹ Department of Statistics, University of Gujrat, Gujrat, Pakistan

² BS (Honours) in Statistics, University of Gujrat, Gujrat, Pakistan

Correspondence: Aysha Saleem, BS (Honours) in Statistics, University of Gujrat, Gujrat, Pakistan. E-mail: aysha_stat22@yahoo.com

Received: August 7, 2012 Accepted: September 26, 2012 Online Published: October 18, 2012

doi:10.5539/jsd.v5n11p15

URL: <http://dx.doi.org/10.5539/jsd.v5n11p15>

Abstract

It is well evidenced that development is a crucial aspect of any country. The demand of a country to be developed, it is necessary to concentrate on other aspects of economy like social and political development rather than just economic growth. In this paper our basic objective is to develop the model to predict the country's development level on the bases of some social, economic and political indicators and also investigate the role of these indicators on development of a country. These indicators are primarily related to economic, health, education and governance. The development of a country is considered as categorical variable, the categories are already defined by United Nations Development Program (UNDP). These categories (highly developed, developed, developing and under developed) are based on Human Development Index (HDI). The data for this study is obtained for 186 countries from World Bank (WB) and UNDP for the year of 2010. Multilayer Perceptron (MLP) Neural Network Model is used for predicting the country's level of development on the basis of economic, health, education and governance indicators, and the relative importance of these indicators in prediction. Our results show that the indicators; health, education and governance have greater effect on countries development level as compare to the economic indicators. From this investigation, it is suggested that developing and under developed countries should also concentrate on the health, education and governance to improve their development level rather than only increasing the economic indicators.

Keywords: development level, Human Development Index, governance, multilayer perceptron

1. Introduction

Development of countries is very important issue that is under discussion from the last some decades (UN, 2012). This problem is one of those which have increased attention from the researcher in the recent past. The word *development* is a dynamic process and verity of definitions on development of a country now exists. Because every discipline that has studied development of country has different opinions and definitions (UN, 2003; Uddin & Joya, 2007; Ngowi, 2009; Sebudubudu, 2010; Nielsen, 2011). Ever since different disciplines consider different indicators to measure development, so there is substantial diversity between the definitions of development described by different institutions and organizations. Therefore development of any country is a very controversial field.

In the past most part of the literature is considering development as a function of economic growth but only economic growth does not assure development in any country. Countries that raise their income but do not focus on other perspective of economy such as political and social aspects; education, health and governance are actually missing important aspects of development. But now in existing literature researchers are also considering the other aspects of economy rather than just economic growth. Many studies had been made on development of countries but there is no any single or compact measure available that can be utilized to correctly determine the level of development for a country. Researchers are still making every effort to find out the solution of this problem. World Bank (WB) and United Nations (UN) are the pioneers of measuring countries with respect to their different perspectives. These organizations also established some measures to determine the level of development of countries. Human Development Index (HDI) is one measure among many attempts.

Education is a very important instrument of development. In developing countries, education is essential in order to achieve a strong economy. As educated people are more likely to have job, earn more and have a respectable position in society. It is also evident that for low-income countries, extension of primary education is the best

investment. For middle income countries, where primary education is established usually, more spend on secondary education that have a greater impact on economic development (UN, 2003). Empirically, in developing countries a year of education raise wages by 10% (Patrinos & Psacharopoulos, 2004). Thus better education leads not only to higher individual income but is also a necessary precondition for development.

Improvement in health facilities may be as important as enhancement in income in relation with development and wellbeing of human. Good health supports human development as well as it also permit inhabitants to be present at work regularly, to be creative and productive at work, and to work for more years. Better Health also can cause to change the population growth rate, which promotes development. More probable health also makes investment in education and economic growth. As healthy parents are likely to be economically more affluences and thus better able to meet the expense of education (or better education). Empirically, high levels of population health and high levels of national income are present together (Bloom & Canning, 2008). Further Millennium Development Goals also addressed to increase health expenditure per inhabitant.

Good governance is a crucial prerequisite for development. It is not a new concept but in the recent past the term governance, specifically good governance, is being increasingly used in the development literature. Because long term developmental outcomes basically can be establish through good governance and it is paramount in promoting higher living standards everywhere. This is a substantial effect that has been examined by many authors like Kaufmann and Kraay (2002), Sebudubudu (2010) and Pradhan and Sanyal (2011). Recently good governance has become conditional for the disbursement of development assistance to less developed nations because fundamental donors and international financial organizations are even more considering the reforms of good governance, for aid and loans, guaranteed (Santiso, 2001; Sharma, 2007). Additionally, foreign investors are also considering good governance for their investment decisions in any country (Fayissa & Nsiah, 2010). Ojima and Iimi (2005) and Sharma (2007) discussed that good governance is not only important for development but it is also a leading feature of any country in determining whether a country has the capability to utilize its resources efficiently to promote growth and reduce poverty.

Economic growth and development of any country is now viewed as a function of many complex and consistent factors. These may include climatic, social, natural, technological, political, and financial factors (Ngowi, 2009). But there is a causality that exists between these socio-economic factors. For example higher income support health by enhancing the hygienic nutrition, better approach to safe water and sanitation and raise the capacity to acquire good quality health care. Conversely, health perhaps is not only an effect but it is the source of high income as well. Thus sustainable economic growth can be attainable only in concordance with other economy sectors such as educational sector, health sector, and governance sector and so forth. Further the Millennium Development Goals are also concentrating on to improve economic development in the poor countries.

Predicting the countries level of development based on these indicators will identify the shortcomings of countries and will facilitate these countries to get betterment in these areas. It will also helpful for the donors, foreign investors and those organizations that provide aid to different countries, in their decision making process where economic features of the country are not satisfactory to discover countries development level. The major contribution in this study is the use of governance factor to predict the country development. Since in the past according to our knowledge there is no such study considering governance factor to predict the development classification of countries. Further a Multilayer Perceptron (MLP) Neural Network is applied to predict the level of development of countries.

After the introduction, literature review of different studies in this context is discussed in section 2. These studies are primarily related to the governance and development factors. Section 3 provides the information about technique employed to fulfil the objectives, source of data and its availability, description of indicators along with their sub-indicators. Section 4 represents the results of analysis and its interpretation and the main the main conclusions that are drawn from this study are given in section 5.

2. Literature Review

Earlier to 1990, development of any county was commonly viewed through the lenses of economic performance (Ranis et al., 2000). Then countries strive to raise their national income as to improve the wellbeing of their people. When time is passed researchers came to know that just economic growth is not a guarantee for the human welfare in any country. Also around about in 90's, governance especially good governance becomes a burning issue among the researchers. Then many studies related to governance, human development and economic growth comes in front. In many such studies it is concluded that development of any country is related to various aspects of economy. Brautigam (1991) stated that political governments of the countries have an important function in the economics performance. To establish and maintain the formal and informal framework

of institutions and organizations, governments use their power and authority, which regulate social and economic interaction.

Governments should be devoted and neutral to serve their people. Then different policy reform or information systems can be used to reduce the fluctuations of economic performance. Much empirical evidence provided which conclude that governance has an effect on different developmental indicators of a country. Such that Kaufmann, Kraay and Lobaton (1999) investigated that governance has a highly significant positive effect on GDP and adult literacy in addition to this there is negative effect of governance on the infant mortality. As a result improved governance will cause to decrease the infant mortality and increase the GDP and adult literacy and this situation is in favour of development. Further evidence also provides that across countries a strong positive correlation between per capita income and the quality of governance exist (Kaufmann & Aart, 2002). Again the result of this strong positive correlation between income and governance supports the development.

In recent times to re-examine the findings of the Kaufmann et al. (1999; 2002) whether good governance matters in development or not. Uddin and Joya (2007) also evident that it is difficult to achieve rapid per capita income or improve the social indicators without improving the good governance. Governance is also play a significant role in pro-poor economic growth due to the strong relation between them. Poverty and income inequality are the dimensions of pro-poor growth so empirical evidence indicate that the good governance can lead to the reduction in both of them (Haq & Zia, 2009). So there is need to prepare and exercise sound policies to improve the governance to attain the Millennium Development Goals that necessitate halve the poverty by 2015. It is also realize that African countries can also achieve the better economic growth by enhancing the governance (Fayissa & Nsiah, 2010). As in case of Botswana, an African country, was one of the poorest country when it obtain independence in 1966. But by adopting the good governance Botswana has turn into the high middle income country in 2000s from the least developed country in 1960s (Sebudubudu, 2010). Botswana offers lessons to other least developed and developing countries that by enhancing the governance, wellbeing of human and development can be achieved.

Besides governance, studies also suggested that indicators related to health and education have significant effect on country's development. The positive relation among health and income is considered as one of the popular relations in any country's development (Bloom & Canning, 2000). This relationship is generally a reflection of the causal association running from income to health. Cole and Neumayer (2006) argued that a key mechanism through which health affects economic growth is via Total Factor Productivity (TFP). Further they find that poor health affects the TFP negatively and this effect is significant and robust over different specification. Anderson and Morrissey (2006) utilize the four different statistical criteria to identify the poor performing countries for a set. They utilize GDP per capita and infant mortality across two periods, 1980-1990 and 1990-2000. Their main findings indicate that few number of countries come into view as poor performers; those that perform poorly in one period, or on one indicator, usually do not perform badly in the other time period or on other indicators. Result for good performer countries is obtained parallel to this.

To make obvious, the relation of health and education with development of a country, some more studies are conducted in the late 2000s. A cross sectional model is developed by Maria-Carmen and Pilar (2007) considering the 39 African countries and it is observe that there is a positive effect of education on economic development and strong positive effect of economic development on health expenditure. Empirical analysis also revealed a positive impact of GDP per capita on health expenditure per inhabitant (Maria-Carmen & Eva, 2007). These relationships proof that education and health, ultimately, play an effective role in countries overall development. An econometric study is carried out by Maria-Carmen (2009) taking into account both; indicators of development and governance by considering the 38 European countries. A compound index based on three groups of indicators is calculated. These groups are educational indicators: public education expenditure per capita and average total years of schooling, life satisfaction and income per capita, including governance indicators: indexes of voice and accounting and government effectiveness. They concluded that the three groups of indicators are highly correlated and existence of different channels to support the voice of nation is very important to enhance the effectiveness of governments and in due course this will promotes development.

As a result from literature review we have come to know that governance has a strong positive impact on developmental indicators such as infant mortality, GDP, adult literacy and pro-poor economic growth (Kaufmann & Aart, 2002; Haq & Zia, 2009). Governance is important to determine whether a country is capable to utilize its resources efficiently to promote economic growth (see Ojima & Iimi, 2005; Sharma, 2007). Besides the governance, health and education is also a leading features of determining the development of any country (Bloom & Canning, 2000; Cole & Neumayer, 2006; Maria-Carmen & Pilar, 2007). In this study a model is developed to predict the level of development of a country through health, educational, economic and

governance indicators. Furthermore we have investigated the relative importance of these indicators with respect to development level of a country.

3. Data and Methodology

3.1 Source of Data

The data for this study has been obtained for 186 countries for the year 2010 from World Bank (WB) and United Nation Development Program (UNDP) on health, education and economic factors and the study variable development level. The data for governance factor are obtained from the World Bank's project of Worldwide Governance Indicators (WGI). Under this project Kaufmann et al. (1999) divided the governance in three major categories and further suggested two indicators in each category as:

I. Political Governance

i Voice and Accountability

ii Political Stability

II. Economic Governance

i Government Effectiveness

ii Regulatory Quality

III. Institutional Dimension of Governance

i Rule of Law

ii Control of Corruption

For each of these governance types, we have used one indicator for each. We have considered voice and accountability, government effectiveness and rule of law for political, economic and institutional dimension of governance respectively. The reason is that if a country focuses on these three indicators then remaining three can be obtained automatically. As voice and accountability indicator to represent the political governance is when peoples of any country are able to select their government then they will not tumble their government. Therefore having voice and accountability certain the political stability in country. Likewise government effectiveness is chosen for economic governance. As when public services are independent from any pressure then continuation of regulatory quality is assured in any economy. Similarly rule of law is elected to indicate the institutional dimension of governance instead of control of corruption. Since existence of rule of law guarantee the control of corruption in the country.

3.2 Data Description

The study variable, development level of a country is already categorized by UNDP. The categories are highly developed, developed, developing and under developed. The detail of four major independent indicators along with their sub-indicators is as follows:

3.2.1 Educational Indicators

i Mean years of schooling (total years) (MYS)

Average number of years of education received by people ages 25 and older, converted from education attainment levels using official durations of each level (UNDP; 2010).

ii Expected years of schooling (total years) (EYS)

Number of years of schooling that a child of school entrance age can expect to receive if existing patterns of age specific enrolment rates keep on throughout the child's life (UNDP; 2010).

3.2.2 Economic Indicators

i Foreign direct investment, net inflows (US\$) (FDI)

Foreign direct investment are the net inflows of investment to attain a lasting management interest (10 percent or more of voting stock) in an enterprise operating in an economy other than that of the investor. It is the sum of reinvestment of earnings, equity capital, other long-term and short-term capital (WB).

ii Exports of goods and services (% of GDP) (EGS)

The value of all goods and other market services provided to the rest of the world. They include the value of insurance, merchandise, transport, freight, travel, license fees, royalties and other services, such as financial,

communication, information, construction, personal, business, and government services. They exclude investment income, transfer payments and compensation of employees (WB).

iii GDP Growth (annual %)

The annual percentage growth rate of GDP at market prices based at constant local currency. GDP is referred as the total market value of goods and services produced in a country in a given period (WB).

3.2.3 Health

i Age dependency ratio (% of working age population) (ADR)

Age dependency ratio is the ratio of dependents, people younger than 15 or older than 64, to the working age population ages in between 15-64. Data revealed as the proportion of dependents per 100 working age population (WB).

ii Life expectancy at birth (total years) (LEAB)

Number of years a newborn infant anticipate to live if existing patterns of age specific mortality rates at the time of birth remain the similar throughout the child's life (UNDP; 2010).

iii Under 5 mortality rate (per 1000 lives birth) (MR)

Under-five mortality rate is the probability per 1,000 that a newborn baby will die before reaching age five, if subject to current age-specific mortality rates (WB).

iv Health expenditure per capita (US\$) (Health Exp.)

Total health expenditure is the sum of public and private health expenditures as a ratio of total population. It covers the stipulation of nutrition activities, health services (preventive and curative), emergency aid designated and family planning activities for health but does not include provision of sanitation and water (WB).

3.2.4 Governance

i Voice and Accountability (VA)

Measuring perceptions of the degree to which a country's citizens are able to contribute in selecting their government, with liberty of expression, independence of association, and a freedom of media (Kaufmann et al., 2009).

ii Government Effectiveness (GoE)

Measure the perceptions of the quality of the civil services, the quality of public services and the extent of its sovereignty from political pressure, the quality of formulating also implementing new policy. Further the believability of the dedication of government to such policies (Kaufmann et al., 2009).

iii Rule of Law (RL)

Measuring perceptions of the extent to which agents have assurance in and follow the rules formulated by the society, and in particular the quality of contract enforcement, property rights, the police, and the courts, as well as the odds of violence and crime (Kaufmann et al., 2009).

These governance indicators are measured in standardized values ranging from -2.5 to +2.5. The high positive value represent good condition of that indicator and vice versa.

3.3 Data Analysis Technique

For the purpose of data analysis we have used MLP neural network model for prediction of development level of a country on the basis of independent indicators.

3.3.1 Neural Network Models (NN)

Neural networks (NN) models, from the past few years, have been used extensively. They are being used over a wide range of different problems, broadly and efficiently. For instance engineering, finance, geology, physics and medicine are some areas in which NN models are used significantly.

3.3.2 Multilayer Perceptron (MLP)

The MLP is an example of an artificial NN and it is the most popular NN architecture that is in use today. It is a feed-forward, supervised learning network and a function of one or more independent variables (also known as inputs or predictors) that minimizes the error of prediction of one or more dependent variables (also called outputs or target variable). Predictors and targets can be a mix of categorical and scale variables.

3.3.3 Architecture of Multilayer Perceptron

An MLP network is composed of units organized in layers and each layer is consisted of nodes. Each unit in a given layer connects to every unit in succeeding layers hence it is a fully connected network. Information flows through each element in an input-output manner. A common example of such network is given in Figure 1. This architecture is consists of an input layer, hidden layer(s) and output layer. The input layer is the first layer and the output layer is the last layer. The layer(s) in-between are called the hidden layers having on restriction on the number of hidden layers. Different functions of these layers are given below.

Input Layer: The function of input layer is to receive the raw information from the external stimuli and forward it to the next layer that is hidden layer.

Hidden Layer: The hidden layer is in between the input layer and output layer where an activation function is employed to manipulate the information obtained from the input layer. This manipulated information is forwarded to the subsequent layer. The subsequent layer can be another hidden layer or output layer. The activation function $\sum_{i=1}^n x_i w_i$ is a general linear combination of weights and inputs.

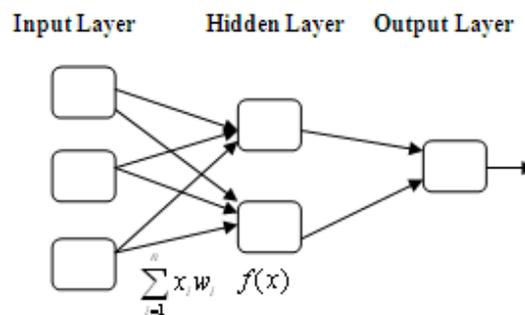


Figure 1. Multilayer perceptron feed forward network

Output Layer: The output layer is the concluding layer to the NN architecture. Function of this layer is same as the hidden layer. This layer also receives the weighted sum of inputs from the hidden layer and utilizes an activation function to produce a value from the network. Usually, each unit in output layer is linked to each unit in the hidden layer. The activation function that implemented to the output layer depends on the measurement level of the output variable.

Activation Function: An activation function is a mathematical conversion of the summarized weighted input units to produce the output. The purpose of an activation function is to link the weighted sums of units in a layer to the values of units in the following layer. Activation functions can be of different types.

- Identity Function:** This function does not have any mathematical formula therefore it looks like $f(x) = x$. As the form of the function shows that it takes real-values and returns them without changed.
- Sigmoid Function:** This function has the form $f(x) = 1/(1 + e^{-x})$. It gets the real-values and converts them into the range of (0, 1).
- Hyperbolic function:** The mathematical form of this function is $f(x) = (e^x - e^{-x}) / (e^x + e^{-x})$. It receives real-values and transforms them into the range of (-1, 1).
- Softmax Function:** This function has the form $\gamma(x_k) = e^{x_k} / \sum_{j=1}^h e^{x_j}$. It takes real-valued arguments and changes them into the range of (0, 1).

4. Results and Discussions

The MLP is applied to predict the development level of a country on the basis of health, education economic condition and governance. It is also investigated that which indicators are playing important role for prediction. Descriptive statistics of all indicators are given in the Table 1. Further analysis is contains MLP network results. Table 2 contains the Case Processing Summary of MLP representing total number of countries available is 186. But MLP consider only those countries for which information on all the indicators is available. For 40 countries information is missing on some indicators hence data on remaining 146 countries are used for analysis. We have used SPSS 16 for applying neural network model on the data it automatically divides the cases into two parts;

one is for training propose and other for testing the model. Data on 99 countries (67.8%) is used for training and for remaining 47 countries (32.2%) are used for testing the model.

Table 1. Descriptive statistics

Indicators	N	Minimum	Maximum	Mean	Std. Deviation
LEAB	186	47.80	83.40	69.62	9.63
MYS	186	1.20	12.60	7.60	2.96
EYS	186	4.40	18.00	12.34	2.96
MR	184	2.10	178.10	42.77	46.52
FDI	183	-4.23E+10	2.36E+11	7.70E+09	2.91E+10
EGS	150	.11	222.96	42.52	29.79
GDP Growth	172	-5.19	15.04	4.16	3.38
ADR	182	16.96	104.80	59.11	17.80
Health Exp.	182	11.90	8361.73	968.97	1681.21
VA	186	-2.17	1.62	-.0795	.99617
GoE	186	-1.74	2.25	-.0361	.98305
RL	186	-1.90	1.97	-.0765	.98159

Table 2. Case processing summary

		N	Percent
Sample	Training	99	67.8%
	Testing	47	32.2%
Valid		146	100.0%
Excluded		40	
Total		186	

Figure 2 is the graphical representation of the MLP neural network. As it is mentioned in pervious section that MLP network is consists of three layers and it can be seen in the given picture. The number of nodes in the first layer (input layer) are equal to the number of independent indicators and bias, so input layer contains the 13 (12 predictors + Bias) nodes. Second layer (hidden layer) contains 5 nodes and in the last layer (output layer) number of nodes are equal to the number of categories in the dependent variable that is 4.

The network automatically chooses hyperbolic tangent activation function for the hidden layer to link the weighted sum of values with the next layer (output layer) and softmax activation function is applied to the output layer to produce the resultant value. In this figure, grey and blue lines show the relationships between input, hidden and output nodes, these relationships are estimated in the form of weights. Grey lines show positive weights and blue lines show negative weights, category 1 (highly developed) has strong positive relation with node 1, 4 and strong negative relation with node 2, 3 of hidden layer. As in the parameter estimation table (table 3) for first category weights for node 1 and 4 are positive and 2, 3 have negative values of the weights.

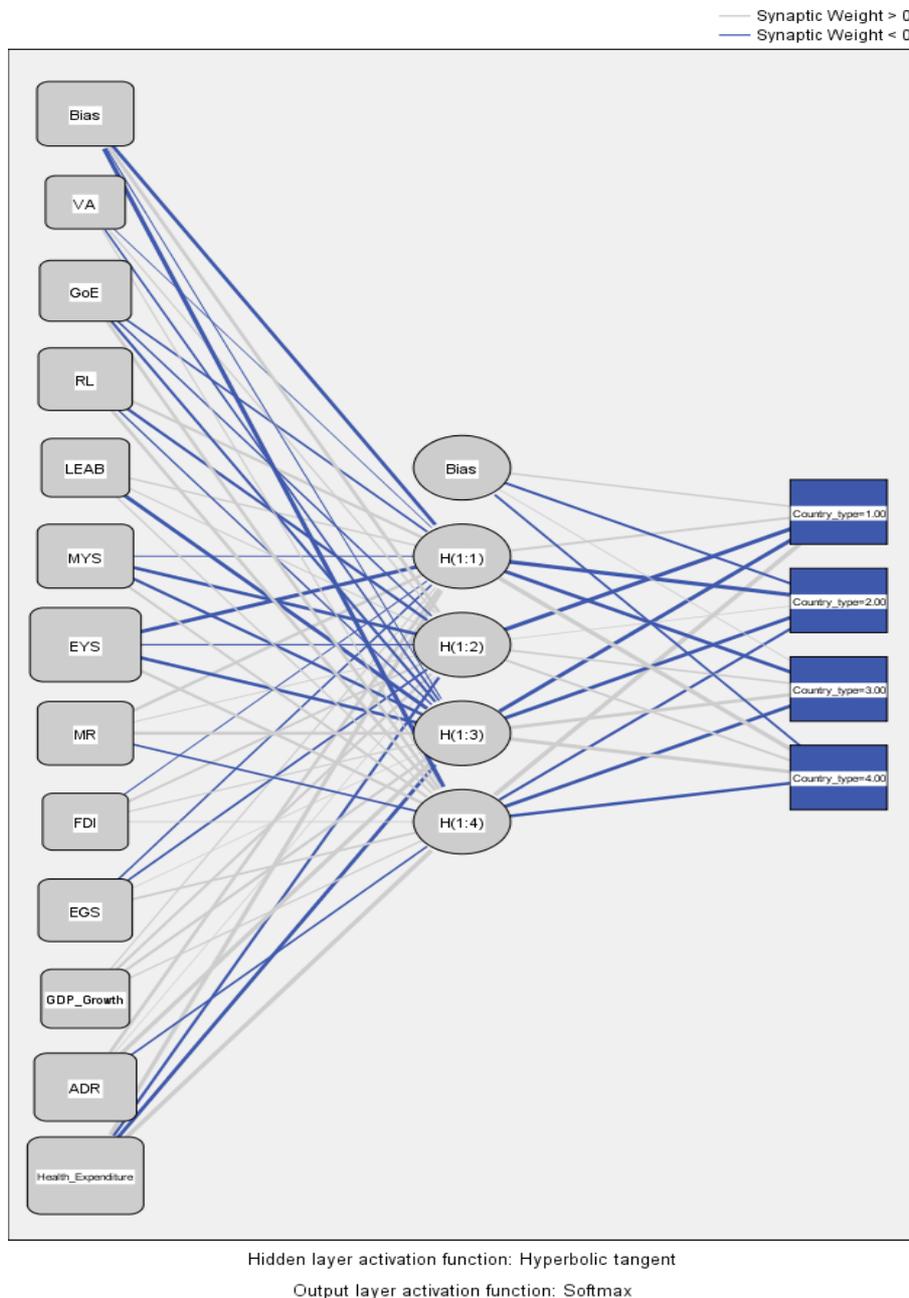


Figure 2. Multilayer perceptron model network

The model summary table (Table 4) provides the information related to the results of training and testing sample. Cross entropy error is given for both training and testing sample. The network minimizes this error function during the training. The small value of this error indicates the better performance of the model regarding prediction. The cross entropy error is less for the testing sample (14.841) as compare to the training sample (30.899). It means that the model performance for testing sample is better than training sample, hence the estimated model can be used for prediction. This table also contains 16.2% and 4.3% incorrect predictions based on training and testing sample respectively, showing the less prediction error for testing sample that indicates good model for prediction. It indicates that the testing sample has small percentage of incorrect prediction as compare to training sample. Table 5 give you an idea about the classification results of the MLP network. In this table cells on the main diagonal of the table are correct predictions and incorrect predictions are given in cells off the diagonal of the table. In the training sample 25 out of 26 countries that are highly developed are classified correctly and one country is misclassified in developed category. 20 out of 25 developed countries are classified correctly and from the remaining countries 2 and 3 countries are misclassified in highly developed and

developing category respectively. Overall, classified correctly training cases are 83.8%, with 16.2% incorrect predictions that is also provided in the model summary table above. Higher percentage of correctly identified cases is a condition for a better model. The testing sample helps to validate the model; classification table shows that overall 95.7% cases in testing sample are correctly classified by the model. So according to the classification table percentage of correct prediction is sufficiently high for training and testing sample. It means the model successfully distinguish among the categories of level of development of the countries.

Table 3. Parameter estimates

Predictor		Predicted							
		Hidden Layer 1				Output Layer			
		H(1:1)	H(1:2)	H(1:3)	H(1:4)	CT=1	CT=2	CT=3	CT=4
Input Layer	Bias	-.899	.675	-.070	-1.236				
	VA	-.034	.098	-.219	.097				
	GoE	-.252	-.170	-.391	.540				
	RL	.465	-.484	-.146	.560				
	LEAB	.175	.081	-1.066	.131				
	MYS	-.059	-.953	-.513	.361				
	EYS	-1.218	-.132	-.835	.493				
	MR	.553	.053	.749	-.262				
	FDI	-.080	.331	.100	.049				
	EGS	-.205	-.277	.016	.319				
	GDP Growth	.097	.400	.444	.123				
	ADR	.592	.015	1.239	-.239				
	Health Exp.	1.120	-.436	-.997	1.345				
	Hidden Layer	(Bias)					.188	-.390	.042
H(1:1)						.358	-1.246	-1.018	2.069
H(1:2)						-1.292	.029	.242	.320
H(1:3)						-1.161	-1.049	.956	1.169
H(1:4)						2.211	-.436	-1.006	-.684

Table 4. Model summary

Training			Testing			
Cross Entropy Error	Percent Incorrect Predictions	Stopping Rule Used	Cross Error	Entropy	Percent Predictions	Incorrect
30.899	16.2%	1 consecutive step(s) with no decrease in error	14.841		4.3%	

The Receiver Operating Characteristics (ROC) curve (Figure 3) is also very helpful for assessing the accuracy of model. The chart has four curves having different colors one for each category of development. The area under the curve is a useful summary statistic to identify the accuracy of any MLP model. The model is less accurate if the under the curve area is close to 0.5. As in this model all the curves are close to upper left corner; which shows that model is best. Table 6 depicts the area under the ROC curves. Area under the curve is close to 1.0 for each category so model is very precise.

Table 5. Classification

Sample	Observed	Predicted				Percent Correct
		Highly Developed	Developed	Developing	Under Developed	
Training	Highly Developed	25	1	0	0	96.2%
	Developed	2	20	3	0	80.0%
	Developing	0	5	16	3	66.7%
	Under Developed	0	0	2	22	91.7%
	Overall Percent	27.3%	26.3%	21.2%	25.3%	83.8%
Testing	Highly Developed	11	1	0	0	91.7%
	Developed	0	14	0	0	100.0%
	Developing	0	1	14	0	93.3%
	Under Developed	0	0	0	6	100.0%
	Overall Percent	23.4%	34.0%	29.8%	12.8%	95.7%

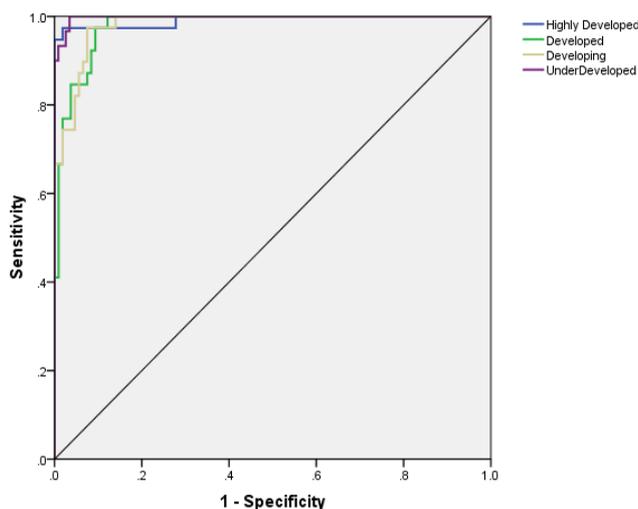


Figure 3. Receiver Operating Characteristics Curve (ROC)

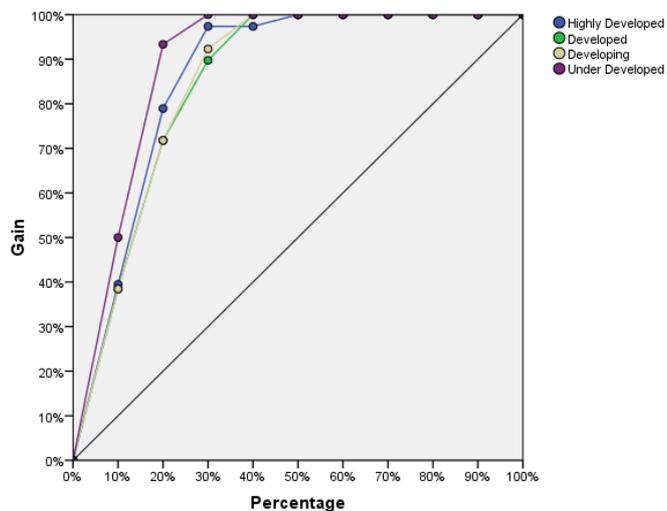


Figure 4. Gain chart

The cumulative gains chart (Figure 4) illustrates the percentage of the overall number of cases in a given category "gained" by targeting a percentage of the total number of cases. For example, the first point on the curve for the highly developed category is at (10%, 40%), means that if all the cases are arranged by predicted pseudo-probability of highly developed category, it would be expecting that the top 10% to have around 40% of all of the cases that actually acquire the category highly developed. Similarly, the top 20% would have more or less 80% of the highly developed. The top 30% of cases would hold about 97% of highly developed, and so forth. The diagonal line is the baseline for these curves; the model will provide gain greater, if curve lies farther above the baseline.

Table 6. Area under the ROC Curve

Category	Country Type			
	Highly Developed	Developed	Developing	Under Developed
Area	0.992	0.979	0.981	0.998

Table 7. Independent variable importance

Indicators	Importance	Normalized Importance
Voice and Accountability	0.011	5.9%
Government Effectiveness	0.066	36.2%
Rule of Law	0.078	42.5%
Life expectancy at birth	0.049	27.0%
Mean years of schooling	0.087	47.8%
Expected years of schooling	0.159	87.0%
Under 5 mortality rate	0.086	46.9%
Foreign direct investment	0.038	21.0%
Exports of goods and services	0.078	42.6%
GDP Growth (annual %)	0.051	28.0%
Age dependency ratio	0.116	63.6%
Health expenditure per capita (current US\$)	0.182	100.0%

The importance of independent variables can be visualized from the Table 7 and Figure 5. The importance of an independent variable is a measure of how much the predicted value of the model changes for diverse values of the predictors. Whereas by dividing the each importance value to the largest importance value gives the normalized importance and then multiplied by 100 to express this in percentages.

The importance chart is basically a bar chart of the values given in the importance table, where bars are sorted in descending value of importance. It appears that indicators of Health (health expenditure per capita (current US\$), age dependency ratio, under 5 mortality rate), Education (expected year of schooling, mean year of schooling), governance (government effectiveness, rule of law) and economic (exports of goods and services) have the larger effect on how the model predict the development level of the countries. Two economic indicators, GDP Growth and Foreign Direct Investment; one indicator from health, life expectance at birth and one from governance, Voice and accountability have the least effect in this prediction of countries development level.

The manual calculation and further detail of MLP model to predict the development level of a country is also given in the Appendix. The purpose of this manual calculation is that by using this model one can predict country's development level when indicators for any near future.

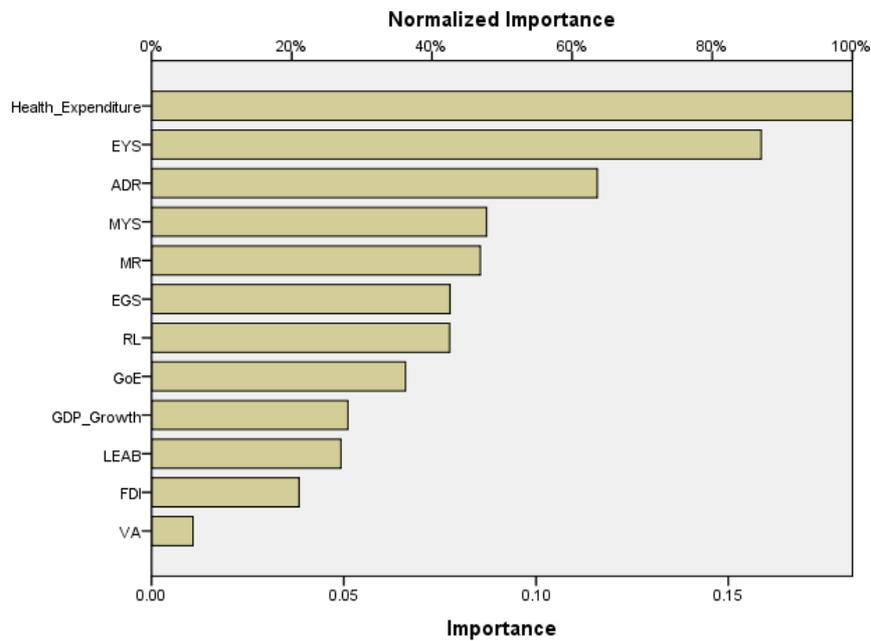


Figure 5. Independent variable importance graph

5. Conclusion

Knowledge of country type with respect to level of development has been very important due to many reasons. So the inspiration for this study is the growing interest of the World Bank, United Nations, governments, international credit unions, private companies, foreign investors and aid donor community about ‘development’ of countries. All of these institutions and people would get advantage if they can determine a country’s development level based on some indicators other than just economic factors. Some countries may have small national income but they can still be considered as developing or developed countries because of their enhancement in other sectors such as health, education and governance. For instance Ukraine and Cuba are considered as developed countries whereas their national income is less than Botswana which is considered as a developing country, just because of their improvement in health and education.

As in this study we develop the MLP model that can be used to predict the development level of countries based on indicators related to health, education, governance and economic factors. The second objective of the study is to examine the relative importance of these indicators that has an effect on the development of a country. It is noted that health indicators that play a more important role are; health expenditure per capita, age dependency ratio and under 5 mortality rate. Both educational indicators; expected year of schooling and mean year of schooling are essential for the development of a country. Whereas two out of three governance indicators that are effective are government effectiveness and rule of law. Only one economic indicator, exports of goods and services, is appearing more effective than the other two indicators. Remaining four indicators that have last four ranking in the variable importance graph are less important as compared to other indicators. Out of these four indicators two are related to economics, one is related to health and one is related to governance. Finally we have achieved our objectives, as the developed MLP model is given in the appendix and through the relative importance of these indicators we can conclude that developing or underdeveloped countries should also concentrate on their health and educational sectors and improve their governance to put the country’s development level upward.

References

- Anderson, E., & Morrissey, O. (2006). A Statistical Approach to Identifying Poorly Performing Countries. *Journal of Development Studies*, 42(3), 469-489. <http://dx.doi.org/10.1080/00220380600576375>
- Bloom, D. E., & Canning, D. (2000). The Health and Wealth of Nations. *Science*, 287(5456), 1207-1209. <http://dx.doi.org/10.1126/science.287.5456.1207>
- Bloom, D. E., & Canning, D. (2008). Population, Health and Economic Growth. Commission on Growth and Development, Working Paper No. 24.

- Cole, M. A., & Neumayer, E. (2006). The Impact of Poor Health on Total Factor Productivity. *Journal of Development Studies*, 42(6), 918-938. <http://dx.doi.org/10.1080/00220380600774681>
- Fayissa, B., & Nsiah, C. (2010). *The Impact of Governance on Economic Growth: Further Evidence for Africa*. Department of Economic and Finance Working Paper Series.
- Haq, R., & Zia, U. (2009). *Does Governance Contribute to Pro-Poor Growth? Evidence from Pakistan*. Pakistan Institute of Development Economics, Working Papers 52.
- Kaufmann, D., Kraay, A., & Lobaton, P. Z. (1999). *Governance Matters*. World Bank Research Working Paper No. 2196.
- Kaufmann, D., & Kraay, A. (2002). *Growth without Governance*. World Bank Policy Research Working Paper No. 2928.
- Kaufmann, D., Kraay, A., & Mastruzzi, M. (2009). *Governance Matters VIII: Aggregate and Individual Governance Indicators 1996-2008*. World Bank Research Working Paper No. 4978.
- Maria-Carmen, G., & Eva, A. (2007). Health Expenditure, Poverty and Economic Development in Latin America, 2000-2005. *International journal of Applied Econometrics and Quantitative studies*, 4(2), 5-24.
- Maria-Carmen, G., & Pilar, E. (2007). Education, Development and Health Expenditure in Africa: A Cross-Section Model of 39 Countries in 2000-2005. *Applied Econometrics and International Development*, 7(2), 135-142.
- Maria-Carmen, G. (2009). Government Effectiveness, Education, Economic Development and Well-Being: Analysis of European Countries in Comparison With the United States and Canada, 2000-2007. *Applied Econometrics and International Development*, 9(1), 39-55.
- Ngowi, H. P. (2009). Economic Development and Change in Tanzania since Independence: The Political Leadership Factor. *African Journal of Political Science and International Relations*, 3(4), 259-267.
- Nielsen, L. (2011). *Classifications of Countries Based on Their Level of Development: How it is done and how it could be done*. IMF Working Paper, 11/31.
- Ojima, Y., & Iimi, A. (2005). *Natural Resources, Economic Growth and Good Governance: An Empirical Note*. JBICI Working Paper No.21. <http://dx.doi.org/10.1016/j.jue.2004.12.007>
- Patrinos, H. A., & Psacharopoulos, G. (2004) Returns to Investment in Education: A Further Update. *Education Economics*, 12(2), 111-134. <http://dx.doi.org/10.1080/0964529042000239140>
- Pradhan, R. P., & Sanyal, G. S. (2011). Good Governance and Human Development: Evidence from Indian States. *Journal of Social and Development Science*, 1(2), 1-8.
- Ranis, G., Stewart, F., & Ramirez, A. (2000). Economic Growth and Human Development. *World Development*, 28(2), 197-219. [http://dx.doi.org/10.1016/S0305-750X\(99\)00131-X](http://dx.doi.org/10.1016/S0305-750X(99)00131-X)
- Santiso, C. (2001). Good governance and Aid Effectiveness: the World Bank and Conditionality. *The Georgetown Public Policy Review*, 7(1), 1-22.
- Sebudubudu, D. (2010). The Impact of Good Governance on Development and Poverty in Africa: Botswana - A Relatively Successful African Initiative. *African Journal of Political Science and International Relations*, 4(7), 249-262. <http://dx.doi.org/10.1080/10220461003763866>
- Sharma, D. S. (2007). Democracy, Good Governance, and Economic Development. *Taiwan Journal of Democracy*, 3(1), 29-62.
- Uddin, M. J., & Joya, L. A. (2007). Development through Good Governance: Lessons for Developing Countries. *Asian Affairs*, 29(3), 1-28.
- UNDP. (2010). *Human Development Report 2010: The Real Wealth of Nation: Pathways to Human Development*. United Nations Development Program. Palgrave Macmillan, New York.
- UN. (2003). *Population, Education and Development: the Concise Report*. United Nations, Department of Economic and Social Affairs, Population Divisions, New York.
- UN. (2012). *Trade and Development Report, 1981-2011: Three Decades of Thinking Development*. United Nations Conference on Trade and Development (UNCTAD), United Nations Publication, New York and Geneva.

Appendix

Prediction Using MLP Model Manually

Prediction of development level of a country using MLP model, the values of independent variables and weights estimated by the model are required. The weights for hidden and output layer are given in Table 3. As the model have three layers; input layer, hidden layer and output layer. So in first step hidden layer receives the weighted sum of incoming signals (information from independent variables) sent by the input layer and processes it in the

hidden layer using Hyperbolic Tangent activation function $f(x) = \frac{e^x - e^{-x}}{e^x + e^{-x}}$. The hidden layer in turn sends an

output signal towards the neurons in output layer. So processing on the weighted information received by the output layer is done on the same way as in hidden layer that is by applying Softmax activation function

$\gamma(x_k) = \frac{e^{x_k}}{\sum_j e^{x_j}}$. This function finally provides the probability for each category of dependent variable. The

country will belong to that category of development level for which this predicted probability is high. Detailed of this prediction processed is give the following, manual calculation for each layer is given below.

For illustration, let us consider the estimation of China's development level. As stated above, for prediction the values of the independent variable are required. In the case of Chine these values are give in the following table. The weights for prediction will be used from Table 3.

Indicators	Values of China's Indicator
Voice and Accountability	-1.650
Government Effectiveness	0.123
Rule of Law	-0.347
Life Expectance at Birth	73.456
Mean Year of Schooling	7.549
Expected Year of Schooling	11.565
Under 5 Mortality Rata	18.400
Foreign Direct Investment	1.85E+11
Exports of Goods of Services	29.572
GDP Growth	10.400
Age Dependence Ratio	38.206
Health Expenditure	220.877

Calculations for Input to Hidden Layer

Let x_i denote the i th independent variable, w_{ij} denotes the weight of i th independent variable and the j th hidden layer node. Let y_j represent the j th node of hidden layer then in our case

$$y_j = bias + \sum_{i=1}^n w_{ij} x_i ;$$

$$i = 1, 2, \dots, n \text{ and } j = 1, 2, 3, 4$$

Now we calculate the results at each node of hidden layer.

For first node

$$y_1 = bias + \sum_{i=1}^{12} w_{i1} x_i$$

$$= -.899 + (-1.650)(-0.034) + (0.123)(-0.252) + (-0.347)(.465) + (73.456)(.175) + (7.549)(-0.059) + (11.565)(-1.218) + (18.400)(0.553) + (1.85081E+11)(-0.080) + (29.572)(-0.205) + (10.400)(.097) + (38.206)(.592) + (220.877)(1.120)$$

$$= -0.899 + (0.056083042) + (-0.030975365) + (-0.161280564) + (12.8548) + (-0.445391) + (-14.08579242) + (10.1752) + (-14806459555) + (-6.062198764) + (1.0088) + (22.61804663) + (247.381842) \\ = -14806459282.$$

For second node

$$y_2 = bias + \sum_{i=1}^{12} w_{i2} x_i \\ = 0.675 + (-1.650)(0.098) + (0.123)(-0.170) + (-0.347)(-0.484) + (73.456)(0.081) + (7.549)(-0.953) + (11.565)(-0.132) + (18.400)(0.053) + (1.85081E+11)(0.331) + (29.572)(-0.277) + (10.400)(0.400) + (38.206)(0.015) + (220.877)(-0.436) \\ = 0.675 + (-0.16165112) + (-0.02089608) + (0.167870523) + (5.949936) + (-7.194197) + (-1.52653908) + (0.9752) + (61261726408) + (-8.19136126) + (4.16) + (0.573092398) + (-96.3022171) \\ = 61261726307$$

For third node

$$y_3 = bias + \sum_{i=1}^{12} w_{i3} x_i \\ = -0.070 + (-1.650)(-0.219) + (0.123)(-0.391) + (-0.347)(-0.146) + (73.456)(-1.066) + (7.549)(-0.513) + (11.565)(-0.835) + (18.400)(0.749) + (1.85081E+11)(0.100) + (29.572)(0.016) + (10.400)(0.444) + (38.206)(1.239) + (220.877)(-0.997) \\ = -0.07 + (0.361240767) + (-0.04806098) + (0.050638629) + (-78.304096) + (-3.872637) + (-9.65651615) + (13.7816) + (18508074444) + (0.473147221) + (4.6176) + (47.33743205) + (-220.214015) \\ = 18508074198$$

For fourth node

$$y_4 = bias + \sum_{i=1}^{12} w_{i4} x_i \\ = -1.236 + (-1.650)(0.097) + (0.123)(0.540) + (-0.347)(0.560) + (73.456)(0.131) + (7.549)(0.361) + (11.565)(0.493) + (18.400)(-0.262) + (1.85081E+11)(0.049) + (29.572)(0.319) + (10.400)(0.123) + (38.206)(-0.239) + (220.877)(1.345) \\ = -1.236 + (-0.1600016) + (0.06637578) + (-0.1942304) + (9.622736) + (2.725189) + (5.70139217) + (-4.8208) + (9068956477) + (9.43337271) + (1.2792) + (-9.1312722) + (297.079087) \\ = 9068956788$$

Now we have to use the activation function tangent hyperbolic to obtain the results of each layer

For first node

$$\tanh(y_1) = \tanh(-14806459282) = -1,$$

For second node

$$\tanh(y_2) = \tanh(61261726307) = 1,$$

For third node

$$\tanh(y_3) = \tanh(18508074198) = 1,$$

For fourth node

$$\tanh(y_4) = \tanh(9068956788) = 1$$

Calculation for Hidden to Output Layer

Let z_k denote the result for k^{th} node of output layer

$$z_k = bias + \sum_{j=1}^4 w_{jk} \tanh(y_j)$$

For first output node

$$\begin{aligned}
z_1 &= bias + \sum_{j=1}^4 w_{j1} \tanh(y_j) \\
&= 0.188 + (0.358)(-1) + (-1.292)(1) + (-1.161)(1) + (2.211)(1) \\
&= 0.188 - 0.358 - 1.292 - 1.161 + 2.211 \\
&= -0.412
\end{aligned}$$

For second output node

$$\begin{aligned}
z_2 &= bias + \sum_{j=1}^4 w_{j2} \tanh(y_j) \\
&= -0.39 + (-1.246)(-1) + (0.029)(1) + (-1.049)(1) + (-0.436)(1) \\
&= -0.39 + 1.246 + 0.029 - 1.049 - 0.436 \\
&= -0.6
\end{aligned}$$

For third output node

$$\begin{aligned}
z_3 &= bias + \sum_{j=1}^4 w_{j3} \tanh(y_j) \\
&= 0.042 + (-1.018)(-1) + (0.242)(1) + (0.956)(1) + (-1.006)(1) \\
&= 0.042 + 1.018 + 0.242 + 0.956 - 1.006 \\
&= 1.252
\end{aligned}$$

For fourth output node

$$\begin{aligned}
z_4 &= bias + \sum_{j=1}^4 w_{j4} \tanh(y_j) \\
&= -0.279 + (2.069)(-1) + (0.32)(1) + (1.169)(1) + (-0.684)(1) \\
&= -0.279 - 2.069 + 0.32 + 1.169 - 0.684 \\
&= -1.543
\end{aligned}$$

Calculation for Output Layer

Finally, for calculating the prediction probabilities for each category of output variable (level of development)

we use Softmax activation function as $\gamma(z_k) = \frac{e^{z_k}}{\sum_j e^{z_j}}$. According to this function

Predicted probability for first category (Highly Developed)

$$\gamma(z_1) = \frac{e^{z_1}}{\sum_j e^{z_j}} = \frac{0.662324276}{4.922205463} = 0.134558437$$

Predicted probability for second category (Developed)

$$\gamma(z_2) = \frac{e^{z_2}}{\sum_j e^{z_j}} = \frac{0.54881636}{4.922205463} = 0.1114971$$

Predicted probability for third category (Developing)

$$\gamma(z_3) = \frac{e^{z_3}}{\sum_j e^{z_j}} = \frac{3.497330629}{4.922205463} = 0.710521057$$

Predicted probability for fourth category (Under Developed)

$$\gamma(z_4) = \frac{e^{z_4}}{\sum_j e^{z_j}} = \frac{0.213738922}{4.922205463} = 0.04342341$$

As a result predicted probability for third category is high as compare to other categories it means; on the bases of considered independent variables China is developing country. According to UNDP 2010 China was declared a developing country hence our model accurately estimate the level of development of China. It is suggested that if data on these indicator is available for the coming years one can easily predict the level of development of the country.