

# **HUMAN DEVELOPMENT INDEX USING IBM WATSON**

## **Mini Project Report**

Submitted By **(BATCH NO: CSE\_014)**

JAKKOJU SHRUTHI (18UK1A05D6)

MUZAFFAR HUSSAIN KHAN (18UK1A0590)

KOKKIRALA AKSHITHA (18UK1A05E6)

RUDROJU SATHISH (18UK1A05A4)

## **INDEX**

<b>CHAPTER 1: INTRODUCTION .....</b>	<b>3</b>
a. <b>OVERVIEW.....</b>	<b>3</b>
b. <b>PURPOSE.....</b>	<b>3</b>
<b>CHAPTER 2: LITERATURE SURVEY.....</b>	<b>3</b>
a. <b>EXISTING PROBLEM.....</b>	<b>3</b>
b. <b>PROPOSED SOLUTION.....</b>	<b>4</b>
<b>CHAPTER 3: THEORETICAL ANALYSIS.....</b>	<b>4</b>
a. <b>BLOCK DIAGRAM.....</b>	<b>4</b>
b. <b>HARDWARE/SOFTWARE DESIGNING.....</b>	<b>5</b>
<b>CHAPTER 4: EXPERIMENTAL INVESTIGATIONS.....</b>	<b>6</b>
<b>CHAPTER 5: FLOW CHART.....</b>	<b>7</b>
<b>CHAPTER 6: RESULTS.....</b>	<b>7</b>
<b>CHAPTER 7: ADVANTAGES AND DISADVANTAGES.....</b>	<b>9</b>
<b>CHAPTER 8: APPLICATION.....</b>	<b>10</b>
<b>CHAPTER 9: CONCLUSION.....</b>	<b>12</b>
<b>CHAPTER 10: FUTURE SCOPE.....</b>	<b>12</b>
<b>CHAPTER 11: BIBLIOGRAPHY.....</b>	<b>12</b>
<b>CHAPTER 12: APPENDIX.....</b>	<b>13</b>

## **CHAPTER 1. INTRODUCTION:**

### **a. OVERVIEW**

The Human Development Index (HDI) is a statistical tool used to measure a country's overall achievement in its social and economic dimensions. The social and economic dimensions of a country are based on the health of people, their level of education attainment and their standard of living. **Calculation of HDI**

- a. Health - Life expectancy at birth
- b. Education - expected years schooling for school-age children and average years of schooling in the adult population
- c. Income - measured by Gross National Income (GNI) per capita

Above three dimensions combined to calculate the Human Development Index (HDI). The value of the Human Development Index (HDI) is between zero and one. Very high, high, medium and low are four main tiers based on the HDI. A country is in the very high tier if its HDI is in the top quartile and the low tier if it's HDI in the bottom quartile.

### **b. PURPOSE**

By Human Development Index Using IBM Watson we will:

- 1. Know fundamental concepts and can work on IBM Watson Studio.
- 2. Gain a broad understanding of Binary Classification.

## **CHAPTER 2. LITERATURE SURVEY:**

### **a. EXISTING PROBLEM**

While exploring this dataset it is clearly visible the relationship between Human Development Index (HDI) with the other dependent variables. (Life Expectancy at Birth, Expected years

schooling for school-age children and average years of schooling in the adult population, Gross National Income (GNI) per capita (PPP US\$)) As shows in the below graphs these dependent variables takes a high value in the countries where Happiness score is a high value.

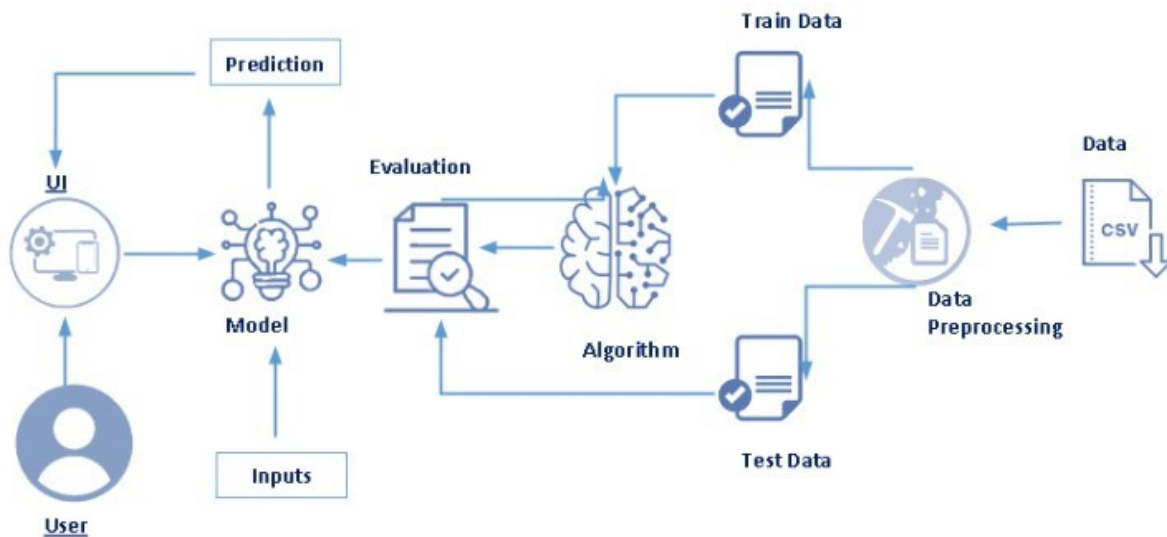
## b. PROPOSED SOLUTION

In the perspective of Machine learning, for a real-world problem like this we can use Linear regression algorithm to predict Human Development. Prediction of Human Development will help the countries to analyze and identify dependent variables and do required steps to develop their Human Development Index (HDI).

## CHAPTER 3. THEORITICAL ANALYSIS:

### a. BLOCK DIAGRAM

#### ARCHITECTURE



b. HARDWARE AND SOFTWARE REQUIREMENT:.

Software Requirement:

REQUIREMENT	SPECIFICATION
Anaconda Navigator	You must have anaconda installed in your device prior to begin.
Spyder, Jupyter Notebook, Flask Framework	1. One should have Spyder and Jupyter notebook. 2. One should install flask framework through anaconda prompt for running their web application 3. We need to build the model using jupyter notebook with all the imported packages.
Web browser	For all Web browsers, the following must be enabled: <ul style="list-style-type: none"><li>• cookies</li><li>• JavaScript</li></ul>

Hardware Requirement:

REQUIREMENT	SPECIFICATION
Operating system	Microsoft Windows UNIX Linux®
Processing	Minimum: 4 CPU cores for one user. For each deployment, a sizing exercise is highly recommended.
RAM	Minimum 8 GB.
Operating system specifications	File descriptor limit set to 8192 on UNIX and Linux
Disk space	A minimum of 7 GB of free space is required to install the software.

## **CHAPTER 4. EXPERIMENTAL INVESTIGATIONS:**

Analysis or the investigation made while working on the solution:

While working on the solution we investigated on what are Health - Life expectancy at birth, Education - expected years schooling for school-age children and average years of schooling in the adult population ,Income - measured by Gross National Income (GNI) per capita, IBM cloud, IBM Watson studio, Machine Learning service, Cloud Object Storage.

The key role on investigation is collection of dataset.

### **IBM Cloud Account:**

IBM Acquired soft layer, a public cloud platform, to serve as the foundation for its IaaS offering. In October 2016, IBM rolled the soft layer brand under its Blue mix brand of PaaS offerings, giving users to access both IaaS and PaaS resources from a single console. IBM cloud provides a full-stack, public cloud platform with various products in the catalog, including options for compute, storage, networking, end to end developer solutions for app development, testing and deployment, security databases, and cloud native services.

Creating the IBM cloud account by going to the IBM cloud login page and click create on IBM cloud account. Enter our IBM id and an ID is created based on the email that we enter. Completing the remaining fields with our information and click create account by this the account is created.

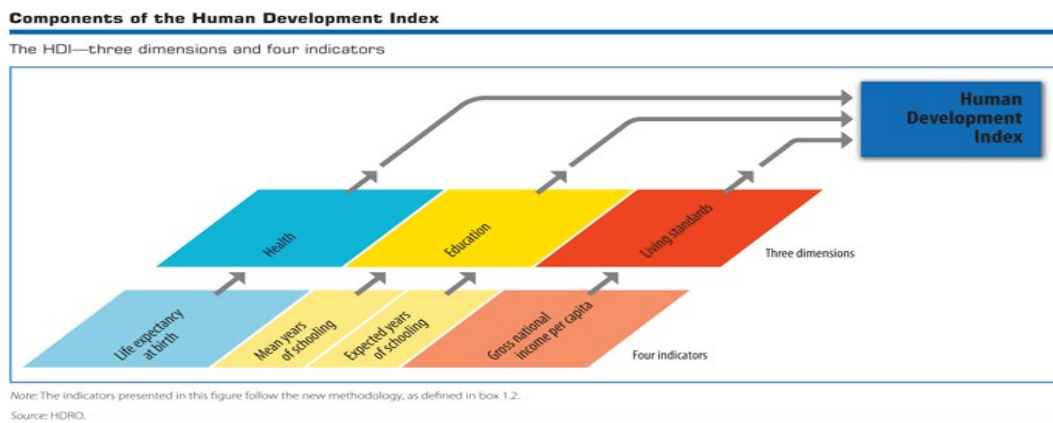
### **Dataset preprocessing:**

The data preprocessing on HDI is done by following steps:

1. Articulate the problem early.
2. Establish data collection.
3. Check our data quickly.
4. Format data to make it consistent.

5. Reduce data.
6. Complete data cleaning.
7. Decompose data.
8. Take the required fields of data

## CHAPTER 5. FLOW CHART:



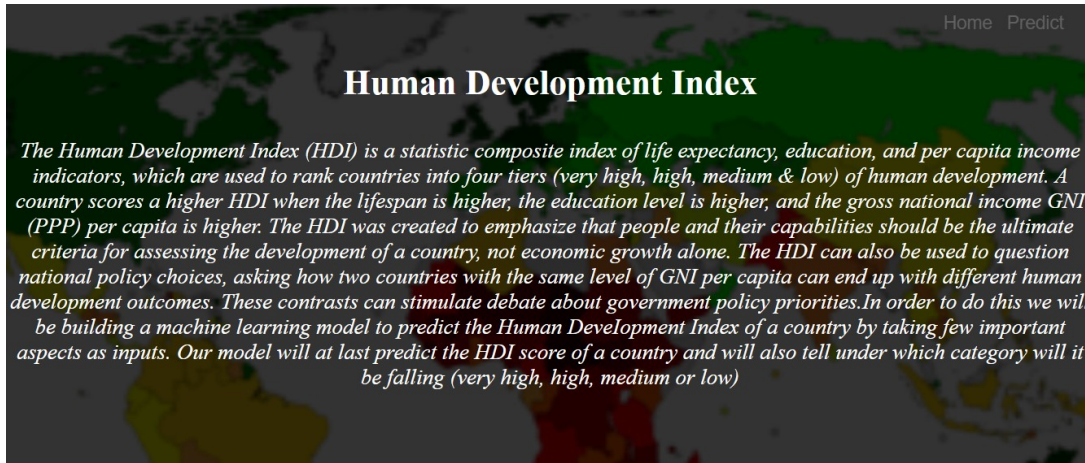
The following diagram shows the control flow of the solution.

## CHAPTER 6. RESULT:

The output is as follows:

Predicting The Results

This is the home page where a brief introduction is given about the Human Development Index and its importance. To see the prediction, you need to click on predict on the top right corner. You will be navigated to the prediction page.



Here in the prediction page, we get to choose the country for which we want to predict the HDI score, enter the values in the required areas in the mentioned ranges. Then click on predict to see the predicted result.

We see that the prediction of the model is High. The country has a Medium HDI score. In the same way we can try different countries.





## CHAPTER 7. ADVANTAGES & DISADVANTAGES:

### Advantages:

- . **Wide use:** HDI indicators are used worldwide. Countries use HDI to compare their level of economic development and the global economic patterns.
- . **Increased infrastructure:** Increase in education level and health of individuals' leads to an improvement in the country's infrastructure.
- . **Balance in human development:** HDI not only concentrates in economic development but also looks at other areas to measure human development like the social measures and individuals' health.
- . **Question national policies:** HDI can enable policymakers in the country to easily adjust and implement the economic policies. It also helps question how countries with the same GNI per capita have different HDI.
- . **Accuracy:** It takes measurement in three areas; health, education and income level making it more accurate.
- . **Reliable:** HDI is more reliable since it involves more than the economic development but also look at the standards of living and level of literacy in measuring the country's development making it more reliable.
- . **Measures per person contribution:** HDI uses GDP per capita measures to determine

the average contribution or gain of each individual in economic welfare and development instead of using GDP.

. **Determines areas that need urgent attention:** HDI data enable the government to know areas that need immediate attention and also to come up with appropriate measures for development.

. **Allocation of funds:** Government can use HDI data to allocate funds in development projects or seek financial aid from the international market to develop underdeveloped areas.

. **Measure the country's status:** HDI measures the social economic development of the country in various aspects.

### **Disadvantages:**

. **Wide divergence among countries:** Different countries have different HDI scores and access different groups differently resulting in wide divergence within the countries.

. **Reflect on long-term changes:** HDI focuses mostly on the long-term changes in the country like the life expectancy of people and has less respond to short-term changes.

. **No clear indication of the country's welfare:** There is no correlation between having a higher national wealth and the welfare of the country. GNI is directly proportioned on how it is spent and may not increase the economic welfare.

. **Measure data on few areas:** HDI doesn't put other factors like gender equality, death rate, poverty and wealth distribution into consideration when measuring economic welfare and development.

. **No standard education index in society:** There is no clear indication of the level of education on all the groups of people in the society. It is difficult to measure whether poor families can access primary, secondary, tertiary and high education in society.

. **Increases inequality:** Sometimes spending more on gross national investments (GNI) per capita can hide the widespread inequality within the country. Higher GNI results in high levels of inequality.

- . **Depends on some factors:** Economic welfare depends on access to clean water, the threat of war, levels of population among others.
- . **Unequal distribution:** There is a lot of criticism that the GDP doesn't measure unequal distribution within the country thus giving inaccurate level of economic development.
- . **Arbitrary measures:** There are various ways of measuring the health or life expectancy and education levels thus, making some of the chosen measures to be arbitrary.
- . **Lack of consistent:** Most countries do not release data required to calculate the HDI on yearly basis thus lacking year to year consistency.

## **CHAPTER 8. APPLICATIONS:**

- Human development refers to the process of widening the choices open to the individual, in the light of the basic needs (UNDP, 2000).
- It is a process to improve people's capability to do or to be what they consider valuable (Bagolin, 2003).
- The HDI is a measurement system used by the United Nations to evaluate the level of individual human development in each country.
- The HDI uses components such as average annual income and educational expectations to rank and compare countries.
- The Human Development Index (HDI) provides a single index measure to capture three key dimensions of human development: a long and healthy life, access to knowledge and a decent standard of living.
- The HDI utilizes four key metrics: life expectancy at birth – to assess a long and healthy life.

## **CHAPTER 9. CONCLUSION:**

In conclusion, we proposed a machine learning model to predict the Human Development Index of the country using the machine learning tools. This system can be used to predict the index of human development of a country and also know into which category it falls into like low, very low, medium, high and very high. From the result, when tested among different attributes like life expectancy, mean years of schooling, GNI per capita and etc in dataset we got accuracy of 95%.

## **CHAPTER 10. FUTURE SCOPE:**

A comprehensive picture of human development calls for composite indices to be complemented by new ways of presenting data such as a dashboard of relevant indicators, to achieve in depth understanding of specific issues. When assessing human progress, the 2016 Report look for comprehensive ways of presenting data, perhaps also including dashboards to measures of gender equality and sustainability.

The 2016 Report will aim to make a contribution to the 2030 Agenda for Sustainable Development. The Report will explore how human development and Agenda 2030 indicators can reinforce one another.

## **CHAPTER 11. BIBLIOGRAPHY:**

[1] "Human Development Index HDI," 26 March 2009. [Online]. Available: <http://wikiprogress.org/articles/initiatives/human-development-index/>.

[2] "Human Developments Reports," UNITED NATIONS DEVELOPMENT PROGRAMME, [Online]. Available: <http://hdr.undp.org/en/data>.

## CHAPTER 12.APPENDIX:

CODE FOR THE SOLUTION WE BUILT:

Source Code:

<https://github.com/smartinternz02/SI-GuidedProject-5809-1634107753/tree/main/HDI>

**Python app.py:**

```
1 import numpy as np
2 import pandas as pd
3 from flask import Flask, render_template, request
4 import pickle
5 app = Flask(__name__)
6 model = pickle.load(open('HDI.pkl','rb'))
7 @app.route('/')
8 def home():
9     return render_template('home.html')
10 @app.route('/Prediction', methods=['POST', 'GET'])
11 def prediction():
```

```

12     return render_template('indexnew.html')

13@app.route('/Home',methods=['POST','GET'])

14def my_home():

15     return render_template('home.html')

16@app.route('/predict',methods=['POST','GET'])

17def predict():

18     input_features = [float(x) for x in
        request.form.values()]

19     features_value = [np.array(input_features)]

20     features_name = ['Country','Life
        expectancy','Mean years of schooling','Gross national
        income (GNI) per capita','Internet Users']

21     df = pd.DataFrame(features_value,
        columns=features_name)

22     #predictions using the loaded model file

23     output = model.predict(df)

24     print(round(output[0][0],2))

25     print(type(output))

26     y_pred = round(output[0][0],2)

27     if(y_pred >= 0.3 and y_pred <= 0.4) :

```

```

28                                     return
    render_template("resultnew.html",prediction_text      =
        'Low HDI' + str(y_pred))

29     elif(y_pred >= 0.4 and y_pred <= 0.7):

30                                     return
    render_template("resultnew.html",prediction_text      =
        'Medium HDI' + str(y_pred))

31     elif(y_pred >= 0.7 and y_pred <= 0.8):

32                                     return
    render_template("resultnew.html",prediction_text      =
        'High HDI' + str(y_pred))

33     elif(y_pred >= 0.8 and y_pred <= 0.94):

34                                     return
    render_template("resultnew.html",prediction_text      =
        'Very High HDI' + str(y_pred))

35     else :

36                                     return
    render_template("resultnew.html",prediction_text      =
        'The given values do not match the range of values of
        the model. Try giving the values in the mentioned
        range'+str(y_pred))

37                                     return    render_template('result.html',

```

```
prediction_text=output)

38

39 if __name__ == '__main__':
40     app.run(debug=True,port=5000)
41
```