**Human Development Index**

**1.INTRODUCTION**

**1.1.Overview:-**

The Human Development Index (HDI) is a statistical composite index of life expectancy, education, and per capita income indicators, which are used to rank countries into four tiers (very high, high, medium & low) of human development. A country scores a higher HDI when the lifespan is higher, the education level is higher, and the gross national income GNI (PPP) per capita is higher. The HDI was created to emphasize that people and their capabilities should be the ultimate criteria for assessing the development of a country, not economic growth alone. The HDI can also be used to question national policy choices, asking how two countries with the same level of GNI per capita can end up with different human development outcomes. In this project  we will be building a machine learning model to predict the Human Development Index of a country by taking a few important aspects as inputs. Our model will at last predict the HDI score of a country and will also tell under which category it falls into  (very high, high, medium or low).

**1.2.Purpose:-**

The purpose of calculating the HDI is to rank global economies by the level of HDI and to compare such a ranking with those that are exclusively based on the GDP per capita .Human development is defined as the process of enlarging people’s freedoms and opportunities and improving their well-being. Human development is about the real freedom ordinary people have to decide who to be, what to do, and how to live.

**2.LITERATURE SURVEY**

**2.1. Existing problem:-**

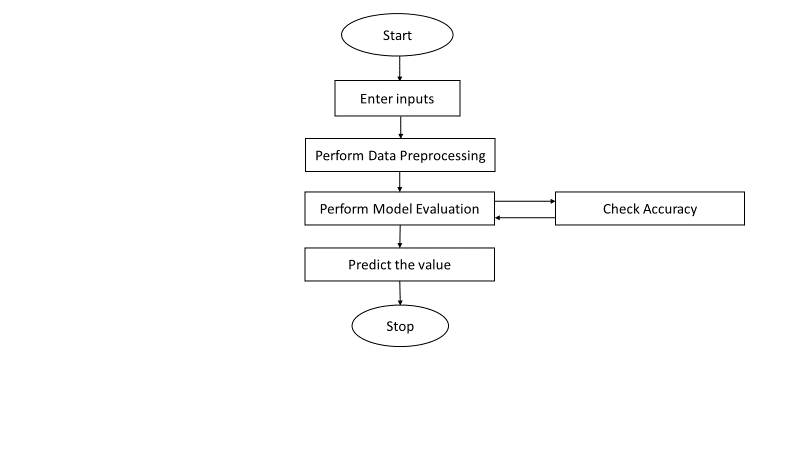
Design and successful operation of Human Development Index is a very complex task and requires the skills of many interdisciplinary skills.The HDI uses components such as average annual income and educational expectations to rank and compare countries.

**2.2. Proposed solution:-**

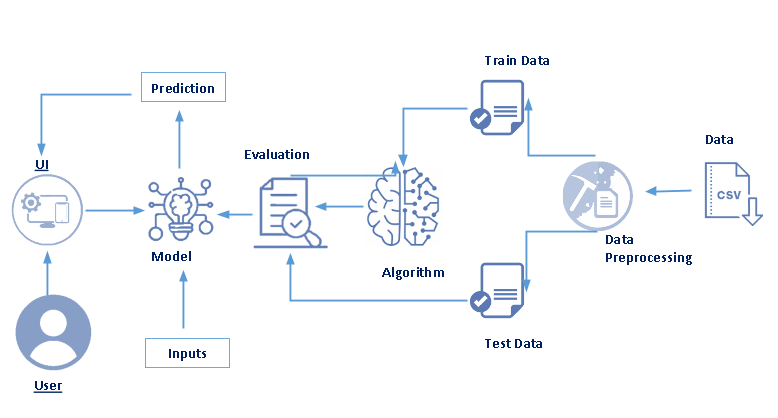
* The proposed solution for the problem is first we need to collect the data related to the Human Development Index of different countries
* According to the collected data we perform different actions for getting the best value

**3.THEORITICAL ANALYSIS**

**3.1.Block diagram:-**



**3.2.Technical Architecture:-**



**3.3.Hardware/software designing :-**

**Software Requirements:**

* OS – Windows XP,7,8,10
* Jupyter Software
* Spyder Software
* Anaconda Command Prompt

**Hardware Components:-**

* Processor – i3
* Hard Disk Storage – 10 GB
* RAM – 1GB

**4.EXPERIMENTAL INVESTIGATIONS**

The main objective of this research is to investigate the Human DevelopmentIndex (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

**5.Results**

The purpose of calculating the HDI is to rank global economies by the level of HDI and to compare such a ranking with those that are exclusively based on the GDP per capita (PPP US $). Three cases are possible:

1. If the HDI rank is close to GDP per capita (PPP US $) ranking, it means there is a harmony between existing resources and development results.

2. If the HDI rank is higher than the GDP per capita (PPP US $) rank, it means that these areas have used their potentials in the best possible way, that is, development policy is in the function of the entire population.

3. If the HDI rank is lower than the GDP per capita (PPP US $) rank, it means that the allocation of resources in the best possible way; that is, their policy of development is not in the function of the entire population, but favors the ruling classes (oil-exporting countries and similar economies based on the exploitation of natural resources and the mono-cultural economy based on them).

The output is generated related to the analysis based on the data collected related to the prediction of HDI of various countries

**6.ADVANTAGES AND DISADVANTAGES**

**Advantages:-**

* It provides us with a much fuller picture of how well developed a country is,
* It shows us that while there is a general correlation between economic and social development, two countries with the same level of economic development may have different levels of social development. See below for examples.
* Some argue that this is a more human centred approach, concerned more with actual human welfare than just mere economics. It gets more to ‘the point’ of economic development.

**Disadvantages:-**

* Relying on the HDI score alone may disguise a lack of social development in a country – for example a very high GNI can compensate for poor life-expectancy, as is the case in the United States.
* It is still only provides a fairly limited indication of social development – only health and education are covered – there are many other ways of measuring health and education.

**7.APPLICATIONS**

* Effective in analysing progress made by countries
* Life expectancy considered -ability to lead a long and healthy life
* Opportunities for education -acess to knowledge and expand choices

**8.CONCLUSION**

The concept of human development had not changed since 1990 when it was also defined in the first Human Development Report. It has remained focused on the lives, freedoms, and abilities of people. The success in the advancement of human development must be seen through the lives of people living and the skills they have. By analyzing the HDI, we conclude that among the 189 countries observed there are significant differences in the level of Life expectancy at birth, Mean Years of Schooling and Gross national income (GNI) per capita. It does not necessarily mean that countries with the maximum value of certain factors constituting the HDI have a higher HDI value. This is because HDI represents the geometric mean of all three elements that together make up HDI. In the period 1990–2017, at the global level, we have positive HDI growth, as a result of positive movements of all elements. As for the countries of the Western Balkans, they are in the group High Human Development and High Human Development, which is not a minor result given the crisis year at the end of the 20th and the beginning of the 21st century.

**9.FUTURE SCOPE**

Human development is the process characterized by the variation of material conditions. These conditions influence the possibilities of satisfying needs and desires. They also explore and realize the physical and psychic, biological and cultural, individual and social potentials of each person. It is also the name of the science that seeks to understand how and why the people of all ages and circumstances change or remain the same over time. It involves studies of the [human condition](https://en.wikipedia.org/wiki/Human_condition) with its core being the [capability approach](https://en.wikipedia.org/wiki/Capability_approach). The [inequality adjusted Human Development Index](https://en.wikipedia.org/wiki/Inequality_adjusted_Human_Development_Index) is used as a way of measuring actual progress in human development by the [United Nations](https://en.wikipedia.org/wiki/United_Nations). It is an alternative approach to a single focus on economic growth, and focused more on [social justice](https://en.wikipedia.org/wiki/Social_justice), as a way of understanding progress.

**10.BIBILOGRAPHY**

* Anand, S and Sen, A (1994) Human Development Index: Methodology and Measurement, Human Development Report 1994.
* Mishra, Srijit and Nathan, Hippu Salk Kristle (2008) On A Class of Human Development Index Measures, WP-2008-020, Indira Gandhi Institute of Development Research, Mumbai.
* Nathan, Hippu Salk Kristle and Mishra, Srijit (2010), Progress in Human Development: Are we On the Right Path? International Journal of Economic Policy in Emerging Economies, Vol.3. No. 3, 199-221.
* Nathan, Hippu Salk Kristle, Mishra, Srijit and Reddy. B. Sudhakara (2008), An Alternative Measure of HDI, WP-2008-001, Indira Gandhi Institute of Development Research, Mumbai

**11.APPENDIX**

**Source Code(Model Building):-**

#Importing the libraries

import pandas as pd

import numpy as np

import matplotlib.pyplot as plt

import seaborn as sns

%matplotlib inline

#Importing the dataset

Development = pd.read\_csv("HDI.csv")

#Listing the first five rows of the dataset

Development.head()

## Listing the first 20 rows for better results of visualization

data1= Development.nlargest(20, "HDI")

data1

Development["Country"].unique()

#Data Exploration

#Country

g = sns.stripplot(x="Country", y='HDI', data=data1, jitter=True)

plt.xticks(rotation=90)

#Data Exploration

#Mean Yearsof Schooling

g = sns.stripplot(x="Mean years of schooling", y="HDI", data=data1, jitter=True)

plt.xticks(rotation=90)

#Data Exploration

#Life Expectancy

g = sns.stripplot(x="Life expectancy", y="HDI", data=data1, jitter=True)

plt.xticks(rotation=90)

#Data Exploration

#Gross national income (GNI) per capita

g = sns.stripplot(x="Gross national income (GNI) per capita", y="HDI", data=data1, jitter=True)

plt.xticks(rotation=90)

#Data Exploration

#GNI per capita rank minus HDI rank

g = sns.stripplot(x="GNI per capita rank minus HDI rank", y="HDI", data=data1, jitter=True)

plt.xticks(rotation=90)

#Building the correalation matrix

heat = Development.iloc[:,[0,1,2,3,4,5,6,7,67]]

sns.heatmap(heat.corr())

Development.shape

from sklearn.preprocessing import LabelEncoder

le = LabelEncoder()

Development['Country']= le.fit\_transform(Development['Country'])

mapping\_dict ={}

category\_col=["Country"]

for col in category\_col:

le\_name\_mapping = dict(zip(le.classes\_,

le.transform(le.classes\_)))

mapping\_dict[col]= le\_name\_mapping

print(mapping\_dict)

X = Development.iloc[:,[2,5,6,7,67]]

X=pd.DataFrame(X)

#Dependant Variable

y = Development.iloc[:,4].values

y=pd.DataFrame(y)

#finding the sum of null values in the selected coulumns

X.isnull().sum()

#replacing the null values with the mean

X.fillna(X.mean(),inplace=True)

X.isnull().sum()

#replacing the null values with the mean

y.fillna(y.mean(),inplace=True)

#Train and test split

from sklearn.model\_selection import train\_test\_split

x\_train, x\_test, y\_train, y\_test = train\_test\_split(X, y, test\_size=0.1, random\_state=42)

#Linear Regression

from sklearn.linear\_model import LinearRegression

reg = LinearRegression().fit(x\_train, y\_train)

#Train and test score

print("Train: ",reg.score(x\_train,y\_train))

print("Test: ",reg.score(x\_test,y\_test))

#predicting and printing the result

y\_pred=reg.predict(x\_test)

print(y\_pred)

#calculating the R squared value

from sklearn.metrics import r2\_score

r2\_score(y\_test, y\_pred)

x\_test

#testing with few values

y\_pred=reg.predict([[13,72.0,5.2,3341.0,14.4]])

print(y\_pred)

#y\_test Values

y\_test

#y\_pred values

y\_pred

#Predicited Y Versus Testing Y

plt.scatter(y\_test,y\_pred)

plt.xlabel('Y Test')

plt.ylabel('Predicted Y')

#saving our model into a file

import pickle

pickle.dump(reg,open('HDI.pkl','wb'))

**Source Code(Application Building):-**

# importing the necessary dependencies

import numpy as np #used for numerical analysis

import pandas as pd # used for data manipulation

from flask import Flask, render\_template, request

# Flask-It is our framework which we are going to use to run/serve our application.

#request-for accessing file which was uploaded by the user on our application.

import pickle

app = Flask(\_\_name\_\_) # initializing a flask app

model = pickle.load(open('HDI.pkl', 'rb')) #loading the model

@app.route('/')# route to display the home page

def home():

return render\_template('home.html') #rendering the home page

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

def prediction():

return render\_template('indexnew.html')

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

def my\_home():

return render\_template('home.html')

@app.route('/predict',methods=['POST'])# route to show the predictions in a web UI

def predict():

#reading the inputs given by the user

input\_features = [float(x) for x in request.form.values()]

features\_value = [np.array(input\_features)]

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

df = pd.DataFrame(features\_value, columns=features\_name)

# predictions using the loaded model file

output = model.predict(df)

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

print(type(output))

y\_pred =round(output[0][0],2)

if(y\_pred >= 0.3 and y\_pred <= 0.4) :

return render\_template("resultnew.html",prediction\_text = 'Low HDI'+ str(y\_pred))

elif(y\_pred >= 0.4 and y\_pred <= 0.7) :

return render\_template("resultnew.html",prediction\_text = 'Medium HDI '+str(y\_pred))

elif(y\_pred >= 0.7 and y\_pred <= 0.8) :

return render\_template("resultnew.html",prediction\_text = 'High HDI'+str(y\_pred))

elif(y\_pred >= 0.8 and y\_pred <= 0.94) :

return render\_template("resultnew.html",prediction\_text = 'Very High HDI'+str(y\_pred))

else :

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 mnetioned range'+str(y\_pred))

# showing the prediction results in a UI# showing the prediction results in a UI

return render\_template('resultnew.html', prediction\_text=output)

if \_\_name\_\_ == '\_\_main\_\_':

# running the app

app.run(debug=True,port=5000)

**12.OUTPUT**

