



# OPEN ACCESS INTERNATIONAL JOURNAL OF SCIENCE & ENGINEERING

## WORLD HAPPINESS ANALYSIS, PREDICATION AND RECOMMENDATION FOR HAPPIER WORLD

<sup>1</sup>Amrat More, <sup>2</sup>Mayuri Muthe, <sup>3</sup>Tabassum Pathan, <sup>4</sup>Prajakta Shelar, <sup>5</sup>C.V.Longani

(<sup>1, 2, 3, 4</sup>) : Students & <sup>5</sup>: Associate Professor Computer Engineering, Sanjivani College of Engineering, Kopergaon, Savitribai Phule Pune University, India.

amratmore51@gmail.com, madhumuthe70@gmail.com, pathantabassum0701@gmail.com, prajaktaishwarshelar@gmail.com, chetnadang16@gmail.com

**Abstract:** The concept of measuring happiness was introduced to the world by the government of Bhutan as Gross National Happiness which later was adapted by UN in the form of happiness index that measures the happiness and well-being of the population of the country. This project aims to analyze the happiness index and its relationship with various factors by making use of different attributes related to happiness of people. Features like ladder score, GDP per capita, life expectancy score etc. are used to analyze the happiness. The results have been presented by using pie charts, graphs and tables. The algorithms and tools like multiple regression and Naive Bayes have been used to measure the extent of dependence of the factors on happiness index. The Happiness Index is framed to set various parameters on grounds of which a country could be ranked in a list of 156 countries. India's rank has come down the list in the year 2019 to be ranked at the 140th position. This clearly indicates India's deteriorating position down the years. This project elaborates the concept of Happiness Index as a measure and analyses various reasons for India to lose its position in the World Happiness Report.

**Keywords:** Data Science, Multiple Linear Regression Algorithm, Naive bayes Classifier.

### I. INTRODUCTION:

We are going to analyze the World and Indian happiness data using Data Science. We will answer the question- Are Indians Happy? We will discuss various parameters that contribute to providing happiness to the citizens like Log GDP per capita scores, Healthy life expectancy scores, Perceptions of corruption scores, Social support scores, Freedom to make life choices scores, and Generosity scores. We will also predict the future values of different parameters related to Happiness of a country. We also propose to generate the recommendations for India to make the country happier, in the ranking of Happiness-Index.

A single factor can have a bigger impact than another. NEF organization in proposes an equation to calculate the happiness index. However, this equation does not consider the economical aspects. Therefore, this work proposes an approach by extending the factors and adopting machine learning techniques to learn about those factors. Due to numerous size of the features, it is unwise to rely on the prediction of a world happiness done by manual analysis. That process will result in high cost of analysis. Therefore,

this work also proposes the use of machine learning to predict the world happiness. Machine learning is a widely known technique to learn about patterns in data. There are several machine learning techniques which can be used to perform a prediction task.

1. Log GDP per capita- It is the value of all the goods and services produced in a country in a given year divided by the total population of the country and adjusted for inflation. This is a widely accepted way of measuring a nation's wealth.
2. Social support- It is the perception of having a supportive social network and having people to depend upon in times of crisis. It raises the quality of life of an individual.
3. Healthy life expectancy- It is the estimate of the expected years of good health life of people at a particular age calculated by the World Health Organization (WHO).
4. Freedom to make life choices- It is the opportunity and liberty to choose and act from at least two options without the influence of any external factor.

5. Generosity- Often regarded as a virtue, it is basically an act of giving or sharing with others or it can be an act of a good deed.

6. Perceptions of corruption- It is a measure of the level of corruption in a country, especially in the public sector

## II. LITERATURE SURVEY:

A Comparative Analysis Of The Factors Affecting Happiness Index Year:-2020 Author:- Parul Oberoi, Shalu Chopra, Yukti Sheth This paper proposes a study that has been carried out to find the reasons behind the country's low happiness index. The factors taken under study for this research are • Social Support • Freedom to make life choices • Physical well-being • Personal safety • Generosity factor that was found to be most influential in determining the happiness of the selected sample The through multiple regression was physical well-being and it was also observed that it was the same factor with the least average i.e., low value of physical well-being among the sample in area of study undertaken is causing low value of happiness index too. [1]

Future Prediction of World Countries Emotions Status to Understand Economic Status using Happiness Index and SVM Kernel. Year:-2019 Author:-B.Prashanthi, Dr. R. Ponnusamy In this paper, a supervised two-tier ensemble approach for predicting a country's BLI score was proposed. The work presented a cost-effective method of BLI prediction with a high degree of efficiency. The capability of the model to predict life satisfaction relied on the proper training features, chosen using a recursive elimination method with 10-fold cross-validation. The work combined three of the top four models, with simple averaging, to enhance the performance of the regression. This is forecasting the Better Life Index (BLI) score using machine learning based regression model that can influence the survival of future generations. 2]

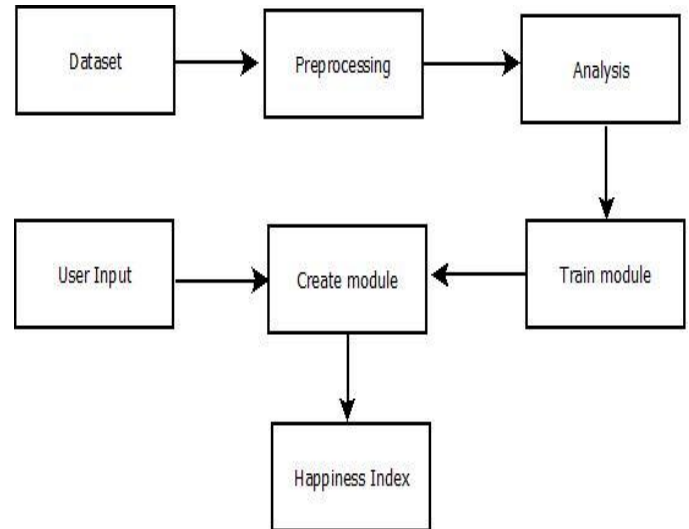
Analyzing Happiness Index as a Measure Along With its Parameters and Strategies for Improving India's Rank in World Happiness Report Year: - 2019. Author: -Sarah Ahtesham This paper elaborates the concept of Happiness Index as a measure and analyses various reasons for India to lose its position in the World Happiness Report. The author appropriately concludes the paper with suitable suggestion. [3]

## III. PROPOSED SYSTEM:

We projected a system which helps finding the happiness index. To implement this various Data science and machine learning techniques have to be used. The model is trained using an appropriate dataset and performance measures. The best model i.e. the model with highest accuracy is used to define Happiness index. we have trained and used algorithms

for prediction. They are Multiple Linear Regression, Naïve Bayes. we have extracted features from preprocessed dataset. we have built all the algorithms and analysis techniques for the analysis and predict Happiness index.

The details about the proposed system is shown in the figure below.



## IV. METHODOLOGY:

### 1. Dataset:

User need to collect real time Happiness data set from authorized organization and then load the data set to the system.

### 2. Pre-processing:

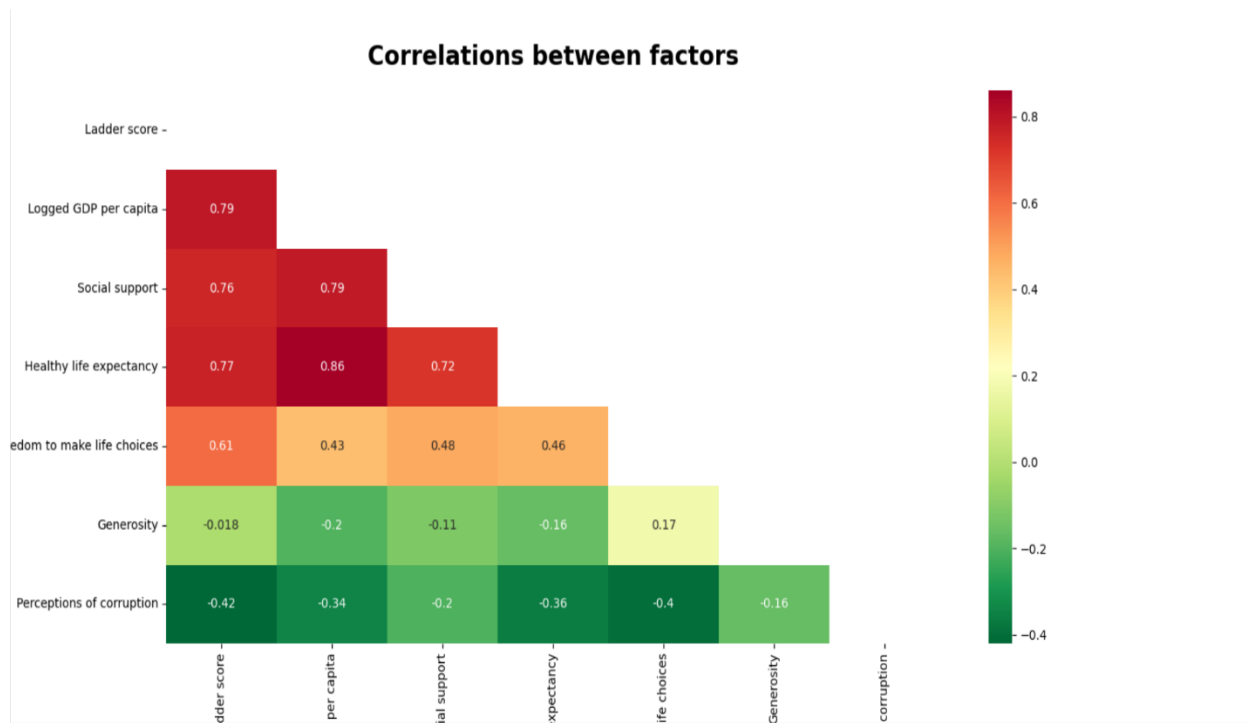
After data split the next system performs preprocessing on the provided dataset. Data preprocessing is a process of preparing the raw data and making it suitable for a machine learning model. Drop the unwanted columns and handle the null values.

### 3. Analysis:

This module will perform the detail analysis and derive many recommendations to the user.

4.1 Correlation:-The relationships between different parameters in the dataset, is analyzed using the method of correlation.

- Perfectly Positive Correlation: When correlation value is exactly 1.
- Positive Correlation: When correlation value falls between 0 to 1.
- No Correlation: When correlation value is 0.
- Negative Correlation: When correlation value falls between -1 to 0.
- Perfectly Negative Correlation: When correlation value is exactly -1.

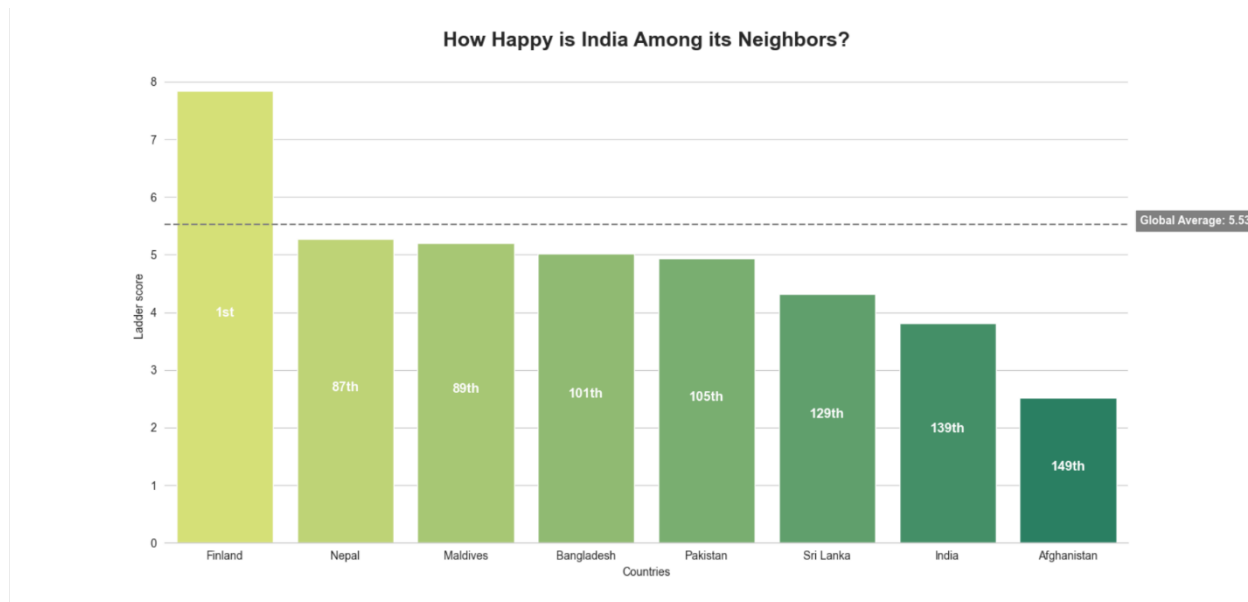


- Positive correlation between logged GDP per capita and ladder score.
- Positive correlation between healthy life expectancy and social support.
- Negative correlation between ladder score and generosity.

4.2 Visualization:- This module display the graphical visualization of analysis and recommendation using the matplotlib & tkinter.library



4.3 Comparison Of Indian With Neighbours:- In this module we have to compare of India with neighbours in the form of graphical representation.



4.3 Recommendation:- Based on the analysis, this module recommends the conclusions based of certain predefined criterion.

#### Training Classifier:

Training classifier takes training data to evaluate the model. Different classifiers were investigated to predict the class of the text. Training is the process of taking content that is known to belong to specified classes and creating a classifier on the basis of that known content. Classification is the process of taking a classifier built with such a training content set and running it on unknown content to determine class membership for the unknown content. There is a lot of classification algorithms like Multiple Linear Regression Algorithm, Naive bayes Classifier.

#### Prediction model:

A prediction model attempts to draw some conclusion from observed values. Given inputs a prediction model will try to predict the value of outcomes. Outcomes are labels that can be applied to a dataset. These model will provide outcome as Happiness Score.

#### User input:

So in the prediction model if the user gives input for different factors values. Once the user gives input to the model it will predict Happiness Score/Ladder Score.

### V. ALGORITHM:

#### 1. Multiple Linear Regression:

Regression Algorithms will be used to predict certain parameters of happiness which are dependent on one or more

other parameters. A multiple regression considers the effect of more than one explanatory variable on some outcome of interest. It evaluates the relative effect of these explanatory, or independent, variables on the dependent variable when holding all the other variables in the model constant.

This regression algorithm has several application across the industry for product pricing, real estate pricing, marketing departments to find out the impact of campaigns.

Multiple regression formula is used in the analysis of relationship between dependent and multiple independent variables and formula is represented by the equation  $Y$  is equal to  $a$  plus  $bX_1$  plus  $cX_2$  plus  $dX_3$  plus  $E$  where  $Y$  is dependent variable,  $X_1$ ,  $X_2$ ,  $X_3$  are independent variables,  $a$  is intercept,  $b$ ,  $c$ ,  $d$  are slopes, and  $E$  is residual value.

$$y = mx_1 + mx_2 + mx_3 + c$$

Where,

$Y$  = the dependent variable of the regression

$M$  = slope of the regression

$X_1$  = first independent variable of the regression

The  $x_2$  = second independent variable of the regression

The  $x_3$  = third independent variable of the regression

$C$  = constant

The Working process of Multiple Linear Regression is as follows:

Step-1: :Data Preprocessing step

Step-2: Fitting Multiple Linear Regression to the Training set

Step-3: Predicting the test result

Step-4: Test accuracy of the result(Creation of Confusion matrix)

Step-5: Visualizing the test set result

## 2.Naive Bayes :

It uses probabilistic approaches and based on Bayes theorem. They deal with probability distribution of variables in the dataset and predicting the response variable of value. An advantage of naïve Bayes classifiers is that they only require less bulk of training data to access the parameters necessary for classification.

This classification technique is based on Bayes theorem, which assumes that the presence of a particular feature in a class is independent of the presence of any other feature. It provides a way for calculating the posterior probability.

$$P\left(\frac{H}{X}\right) = \frac{P\left(\frac{X}{H}\right) \cdot P(H)}{P(X)}$$

$P(H|x)$ = posterior probability of class given predictor  $P(H)$ = prior probability of class

$P(x|H)$ = likelihood (probability of predictor given class)  $P(x)$  = prior probability of predictor

The Working process of Naive Bayes is as follows:

Step-1: Data Pre-processing step

Step-2: Fitting Naive Bayes to the Training set

Step-3: Predicting the test result

Step-4: Test accuracy of the result(Creation of Confusion matrix)

Step-5: Visualizing the test set result.

## VI RESULT:

The objective of this claim is to advance a system which knows about Happiness Index prediction shows user appropriate results with proper accuracy. For that user need to give data set in csv file to the system as input. After that system will done all processing and apply machine learning algorithm to the given input. The dataset covers the information of Happiness Index information. The correctness of testing is 80%. Hence we can say that if a user feeds a particular Happiness analysis in this model, there are 80% chances that it will be prediction is more accurate.

## VII CONCLUSION:

In this research, We have demonstrated the performance of our proposed model for World Happiness Analysis

, Prediction and Recommendation For Happier World. A system has been positively applied for predict the happiness value using machine learning algorithms. Machine learning approach is best to use. System absolutely takes input appearance effort from the user and conveys output in the form of numerical and Graphical format with proper accuracy. The model is trained using an appropriate dataset and performance evaluation is also done using various performance measures. Overall, the use of machine learning algorithms finding the Happiness index. The user will be able to upload data and analyze the different aspects of World Happiness. Also there will be features to draw the conclusions and thus recommendations from the system, which will be very useful to take strategic decisions for happier country. The system can also predict the future values of important parameters related to happiness.

## REFERENCE:

- [1] Parul Oberoi, Shalu Chopra, Yukti Seth, "A Comparative Analysis Of The Factors Affecting Happiness Index", International Journal of Scientific & Technology Research Volume 9, Issue 03, March 2020.
- [2] B. Prashanthi, Dr. R. Ponnusamy, "Future Prediction of World Countries Emotions Status to Understand Economic Status using Happiness Index and SVM Kernel", International Research Journal of Engineering and Technology (IRJET), Volume: 06 Issue: 11, Nov 2019.
- [3] Sarah Ahtesham, "Analyzing Happiness Index as a Measure Along With its Parameters and Strategies for Improving India's Rank In World Happiness", 10.21917/ijms.2020.0161.
- [4] Charles Alba, "A Data Analysis of the World Happiness Index and Its Relation to the North-South Divide", Undergraduate Economic Review, Volume 16, Issue 1, Article 6, 2019.
- [5] Rajesh K Pillania, India Happiness Report 2020, Retrieved from <http://pillania.org/wp-content/uploads/2020/09>.
- [6] World Happiness Report 2021 - Amazon AWS, <https://happiness-report.s3.amazonaws.com>
- [7] World Happiness Report – ResearchGate, <https://www.researchgate.net/publication>.