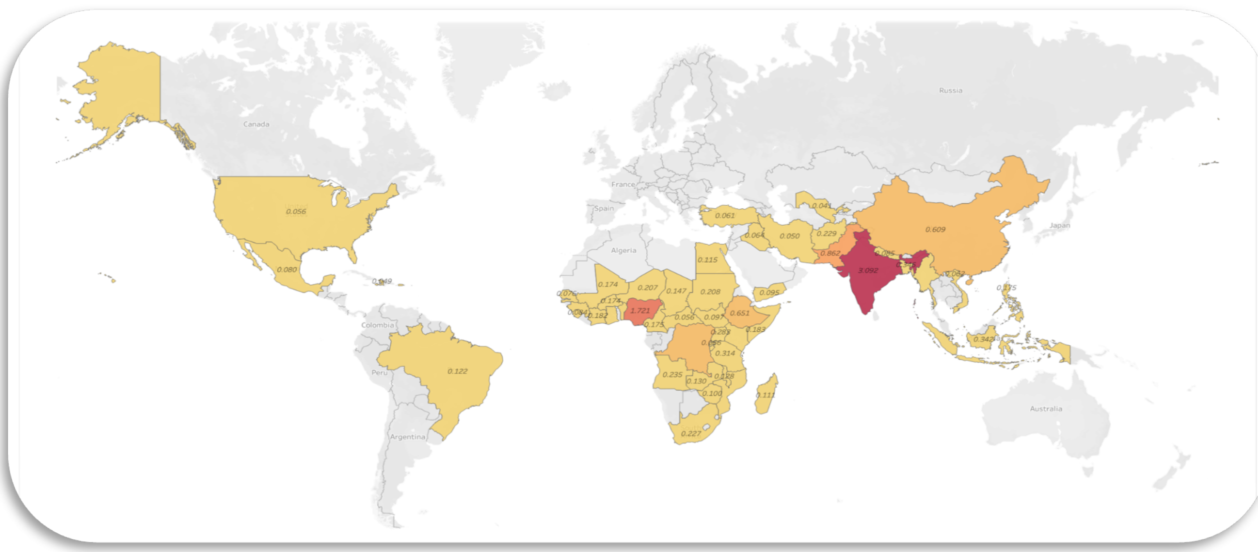
ITWS 6400 – XINFORMATICS

CMIS

Child Mortality Information System



SUBMITTED TO:

Professor
Thilanka Munasinghe

TEAM:

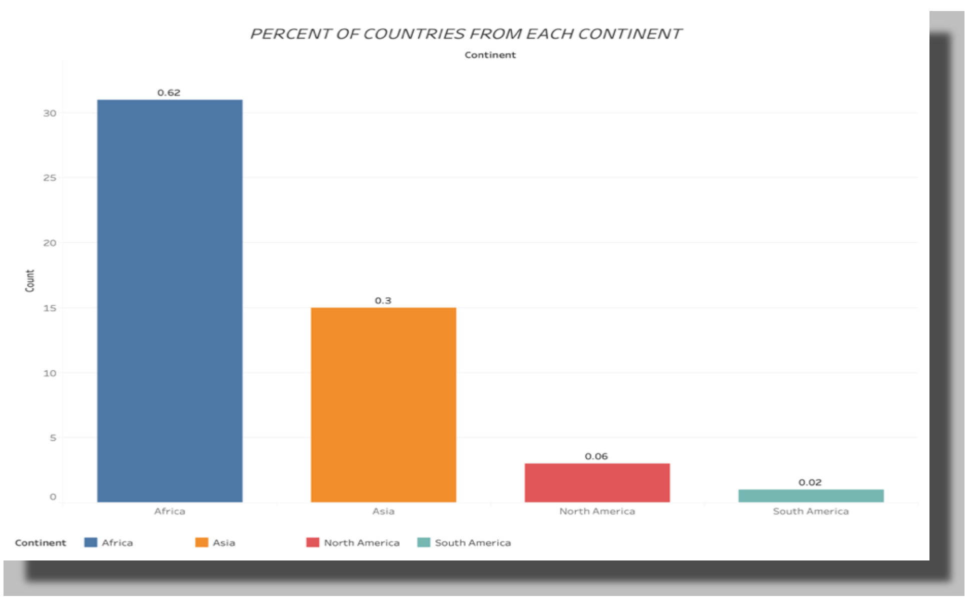
Chirag Sahni
Deepanshu Dey
Himanshu Dey
Shlok Shah
Mukesh Mohanty

TABLE OF CONTENTS

1.1 ABSTRACT

With increase in the figures for overall population, rate of infectious diseases mortality, rate of burden diseases mortality, rate of maternal mortality and increase in average median age across the globe, child mortality ratios for countries have been an important factor of consideration for United Nations and governing bodies. Why is this more important than other factors? The Global HumanPopulation growth amounts to 83 million annually, nearly 1.1% per year.[1] Moreover, average median age, single most important factor for age distribution is also increasing. As per, world data Japan tops the list by 45.9 years with 45 other countries and dependent territories over 40 years – 31 belongs to Europe, 9 to North America and 6 to Asia.[2] With active increase in global infections due to recent novel COVID-19 pandemic outbreak and other life-threatening epidemics before it, such as HIV and AIDS have disrupted quality of life index across countries. Further, "healthcare disruptions could reverse decades of improvements. Hundreds of thousands of additional under-5 deaths may be expected in 2020" are directly referenced United Nations Department of Economic and Social Affairs Sustainable Goals.[3] Hence, child mortality is considered in *UN Sustainable development goal (SDG Goal 3: Good health and Well-being)*.

The overall study scopes different factors such as Gross Domestic Product, Life Expectancy Rate, HealthCare Ex penditure Index, Education Index and Average Female Education Years across countries with highest child morta lity rates. It measures whether above factors have positive, negative or neutral correlation with the mortality rates ? The mortality rate is calculated on the basis of deaths for 'HIV/AIDS', 'Diarrhoeal diseases', 'Pertussis', 'Tetanus', 'Measles', 'Meningitis/Encephalitis', 'Malaria', 'Acute Lower Respiratory Infections', 'Prematurity', 'Birth Asphyxia and Birth Trauma','Sepsis and Other Infectious Conditions of the Newborn', 'Other Communicable, Perinatal an d Nutritional Conditions','Congenital Anomalies', 'Other noncommunicable diseases', and'Injuries'. The specific a ge group considered for the analysis resides between 0-5 Years of the cases, inclusive of 0-27 days and 1-59 mont hs of data. Below is percentage wise distribution of selected countries from continents for mortality rate analysis. Above factors are determined for these specific countries and regression model are applied to provide the predicti ve analysis of child mortality (in terms of deaths) for the target variable. The overall goal is to reduce the mortalit y rates with predictive analysis by 10% annually.

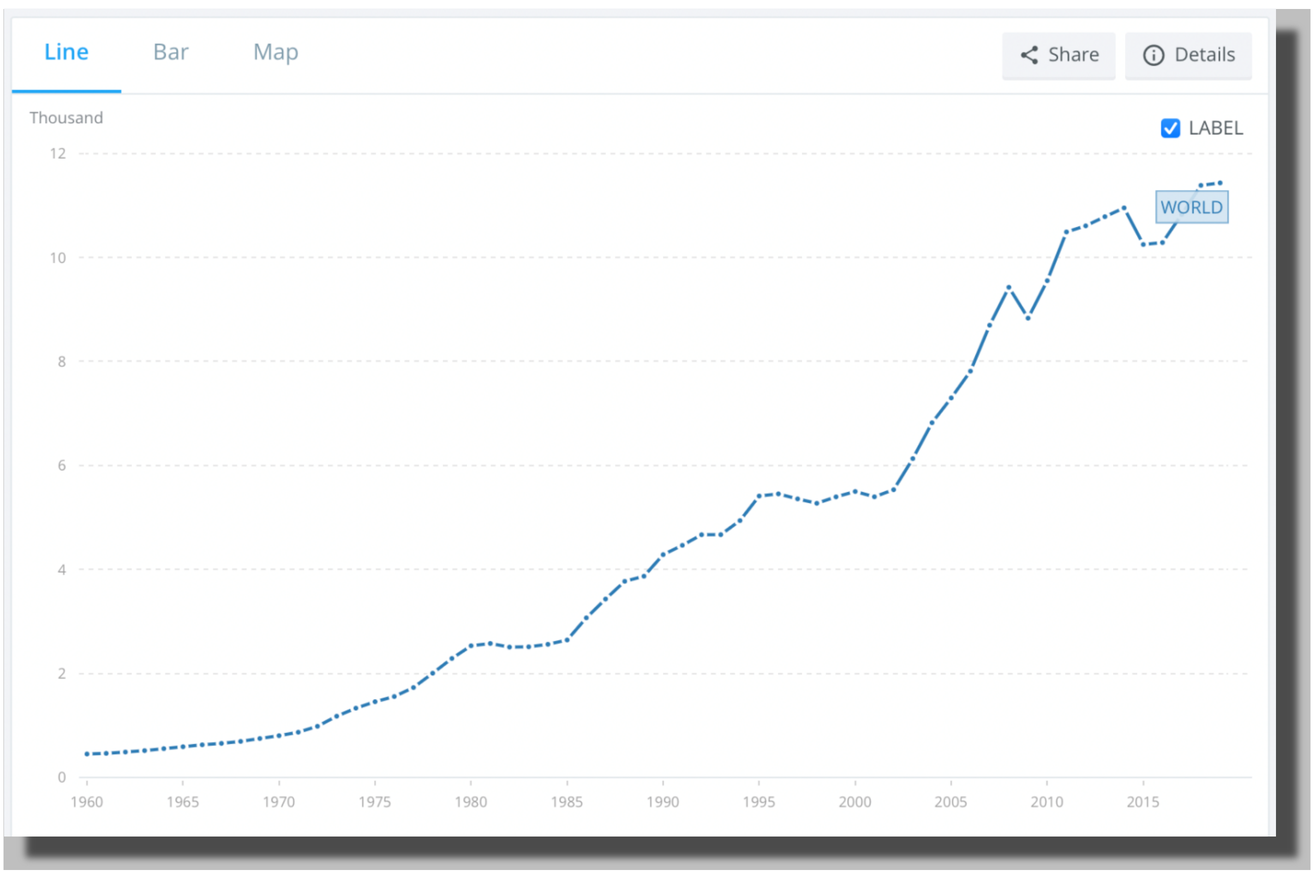


Screenshot from the frontend-app

1.2 INTRODUCTION

"Gross domestic product (GDP) is the total monetary or market value of all the finished goods and services produced within a country's borders in a specific time period. As a broad measure of overall domestic production, it functions as a comprehensive scorecard of a given country's economic health." [4] GDP plays a significant role in determining the mortality factor for a country by directly or indirectly contributing to it. In developing countries, many studies have concluded that people with higher income and higher standard of living have better health rates, increased life expectancy and lower comorbidity and mortality (see Gwatkin et al. 2007 for a review). Although, GDP is not the only factor that directly correlates to the rate of mortality other factors such as healthcare infrastructure, healthcare finance, logistics etc. contributes to it.

By studying the GDP and its growth in terms of child mortality for specific countries and implementing machine learning models we can astutely make predictions for number of deaths. The study analyses and evaluates 15 years of available GDP data for specific countries – having highest deaths for the age groups between 0 to 5 years. However, after model development we can increase the scope to rest of the countries and even train the model to incorporate children age groups of 5 to 10 years and 11 to 15 years.



World GDP Line Graph from 1960 to 2015

Source: https://data.worldbank.org/indicator/NY.GDP.PCAP.CD?end=2019&start=1960&view=chart&year=1960

2. THE GENERICS

2.1 Use Case Development

Use Case: CMIS (Child Mortality Information System)

Point of Contact:

- Chirag Sahni <sahnic@rpi.edu>
- Deepanshu Dey <deyd@rpi.edu>
- Himanshu Dey <deyh@rpi.edu>
- Mukesh Mohanty <mmohan@rpi.edu>
- Shlok Shah <shahs14@rpi.edu>

Version: 1.0

Date: 20th April 2021

Use Case Name

CMIS (Child Mortality Information System)

Goal

The primary goal of the CMIS is to predict the Child Mortality Rate for the top 50 nations with the highest current child mortality index and help in reducing it in the nations across the globes by designing a highly scalable and Health Information System.

This system mainly focuses on collecting information from various related indexes and then processing that information to help the existing governmental health information systems to take necessary measures to placate the impact of numerous diseases on the child mortality.
It further scopes the potential health systems to implement a highly scalable and affordable information system for the end users, thereby reducing the number of economic fatalities.

Summary

The system sets to present a correlation and prediction between the child mortality rate of different nations based on various parameters such as:

- Health Expenditure
- Life Expectancy Index
- Education Index
- Average Years of Schooling for Females
- Gross Domestic Product

Different datasets are collected for the above-mentioned parameters from the respective data sources and analyzed for one-to-one correlations with the child mortality rate for multiple nations.
This system is essentially a web application which encompasses various data visualization techniques to showcase responsive charts to the user based on the parameters selected by the user on the UI.

Under the hood, the system employs various data pre-processing techniques and exploratory data analysis to fuse data from various sources in order to apply suitable machine learning models to predict the child mortality rate for the respective countries for future years, given the estimated values of the above parameters for that time period.

Actors

Primary Actors:

- Child Mortality Data Manager
- Data Source Manager
- Data Administrator
- Principal Administrator
- Users

Secondary Actors:

- World Health Organization
- World Bank
- United Nations
- Various Governmental Health Organizations

Preconditions

- Availability of UN Population Division Data for Life Expectancy.
- Availability of WHO Data for Child Mortality
- Availability of WHO Data for Health Expenditures per capita provided by the World Bank
- Availability of UN Human Development Center Data for Education
- Availability of UN Human Development Center Data for Female Education Index

Post Conditions

- Successful identification of affordable healthcare plans and policies for end users by the governmental health organizations to decrease the annual child mortality rates.
- Optimized distribution of national GDPs to health care sector.
- Designing customer oriented public policies.
- Developing affordable Insurance plans and policies by IOM, AMA etc.

Triggers

- Non-Scalable and Expensive Healthcare Systems as defined by both public and private entities.
- Higher relative mortality rates at institutional or global levels as defined by United Nations, Pervasive Developmental Disorders, Interagency Group for Child Mortality Estimation and Institute for Health Metrics and Evaluation. [4]

Normal Flow

1. Data collection of previous years from different sources (such as WHO, UN, World Bank etc.) by the Data Source Manager.
2. Child Mortality data collection for different nations by the Mortality Data Manager.
3. Implementation of data pre-processing techniques and exploratory data analysis to fuse data from various sources.
4. Data Administrator analyzes this pre-processed data and prepares it in a particular format required for the following steps.
5. After transforming, the data is fed to the 'under the hood' models for predictive analysis of the future child mortality rates.

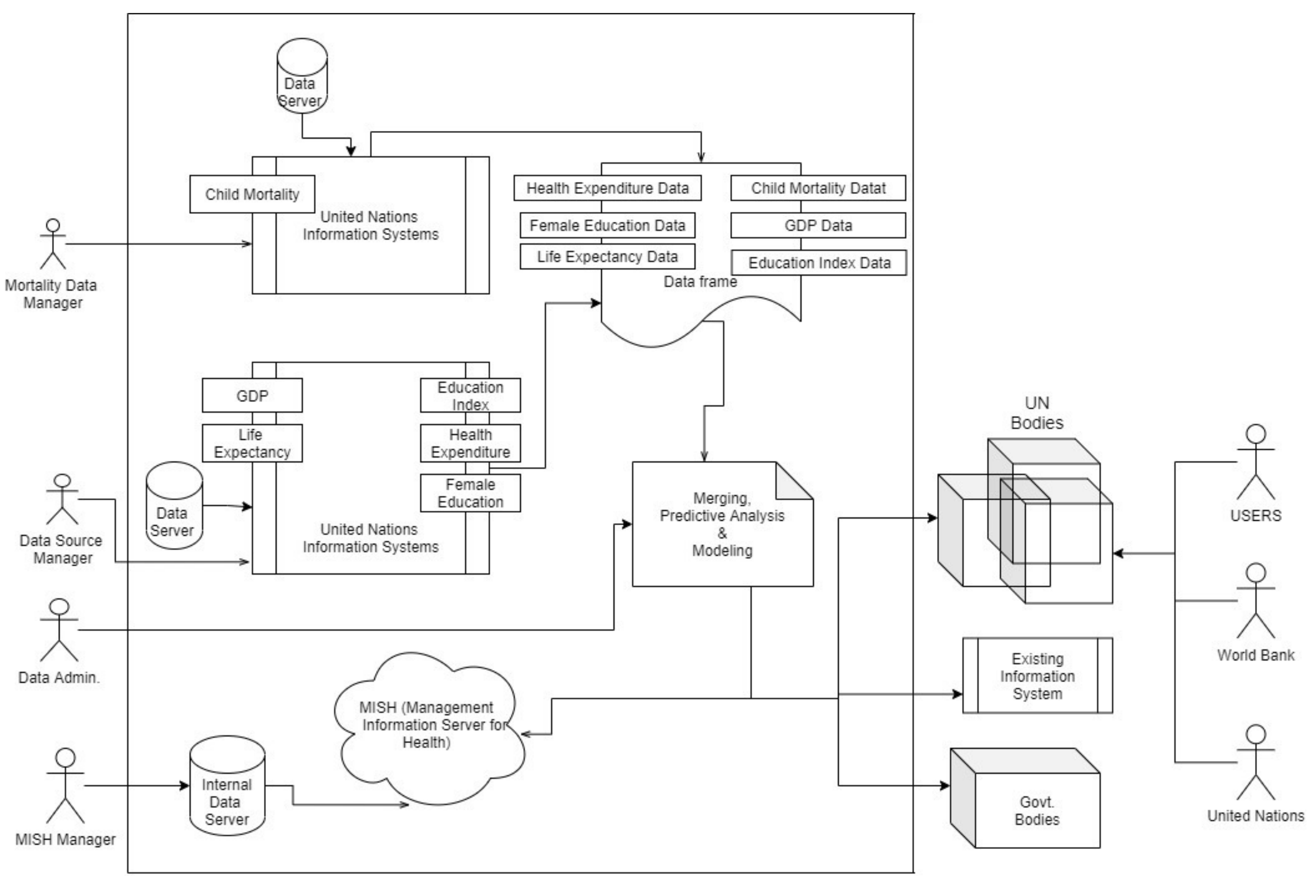
6. Data Administrator feeds the health system data into the Management Information Server for Health (MISH), Cloud based for future reference/research and baseline comparisons.
7. The visualizations developed on the different datasets are portrayed on the web application based on the user selection of parameters such as HCE, LE etc. to showcase the respective responsive charts.
8. Principal Administrator propagates the baseline documents and reports all the details to various national health governing bodies to help develop healthcare plans and policies for end usersto decrease the annual child mortality rates.

Alternate Flows - NA

Exception Flows

- Unavailability of any of the five datasets under consideration would halt the flow of this information system.
- Change in the dataset structure for any of the datasets would require restructuring and refactoring of existing models implemented in the system.
- HHS/FDA non approval for the implementation and restructuring of the existing healthcare systems.
- Limited in cash flows for the design and development of the systems.
- The server hosting the CMIS web application crashes due to any internal / external issues.

Activity Diagram



Includes

None.

Priority

Medium Priority, research and product data can be used by different existing national health systems to control the child mortality rates for multiple nations.

Source

Inefficient health-care information systems providing validated information on multiple factors leading to increasing child mortality rate in developed and developing nations. [5]

Frequency of Use

This information system can be used on annual basis or even twice a year with the baseline checklist (previous year's data).

Business Rules

- Only the Administrator can designate the statistical reports and data from the UN, WHO, World Bank Sources to the agencies for planning and system development.
- Designated database personnel should be entrusted to manage and maintain the data servers.

Special Requirements

- Principal Administrator should be IIPMR Certified Research Professional (CRP) or IIMRA Certified Professional.
- Data Source Manager and Child Mortality Data Manager should be authorized to collect datasets from UN, WHO, IGME, IHME servers.

Assumptions

- Predictor Algorithms used will precisely retrieve the data for top performers/contributors.
- Administrator efficiently assess the product data and plan provided by the agencies or groups.
- The datasets are not corrupted and are also updated regularly to maintain the efficiency of the information system.

Notes

CMIS: Child Mortality Information System
UN: United Nations
WHO: World Health Organization
GDP: Gross Domestic Product
AMA: American Medical Association.
IOM: International Organization of Migration
MISH: Management Information Server for Health
HCE: Health Care Expenditure
LE: Life Expectancy
HHS: Unites States Department of Health and Human Services
FDA: Food and Drug Administration
UN IGME: United Nation Inter-Agency Group for Child Mortality Estimation.
UN IHME: United Nation Institute for Health Metrics and Evaluation.
IIPMR: International Institute for Procurement and Market Research

Resources

In order to support the capabilities described in this Use Case, a set of resources must be available and/or configured. These resources include data and services, and the systems that offer them. This section will call out examples of these resources.

Data:

Data	Type	Characteristics	Description	Owner	Source System
Life Expectancy, total(years)	Protected / Authorized.	Global Servers / Authorized Cloud Servers	Provides the information regarding the total life expectancy in years for all the nations across the world for over 50 years	World Bank	Authorized Data Servers.
GDP Growth	Protected / Authorized.	Global Servers / Authorized Cloud Servers	Provides the GDP Growth data of the countries in US Dollars	World Bank	Authorized Data Servers.
Education Index	Protected / Authorized.	Global Servers / Authorized Cloud Servers	Provides statistical info regarding the education and female education with over 150 indicators	United Nations, WHO	Authorized Data Servers.
Health Expenditure	Open Source	Public Cloud Servers	Provides information of Health Expenditure per Capita data for around 190 countries in the world for the time period of 2000 to 2018	World Bank, WHO	Authorized Data Servers.

Application Services (3)

Application	Owner	Description	Source System
Various National Health Agencies	National Governments	Examples for USA: https://www.cdc.gov https://www.nih.gov https://en.wikipedia.org/wiki/American_Medical_Association	Global Health System of Disease Control and Prevention and similar departments for different nations.

2.2 *Prototype Implementation Design of CMIS*

The designed prototype inculcates all the primary factors such as Gross Domestic Product, Life Expectancy Rate, Health Index, Education Index and Average Female Education Years that are responsible for affecting the child mortality in its development. The design incorporates microservices approach to collect, render and estimate data (deaths) for the given set of parameters. Overall, the architecture modularizes the frontend user-interface and backend microservices with an object-oriented approach to define the client and server. The services (API's) are used to bridge the connection between internal modules and the database (currently static). Child Mortality Information Systemis developed using the component modelling framework, ***react*** – an open-source, front end, Java Script library for binding the user interface or UI components.[6] The interactive UI design helps the user to visualize the graphical correlations for the above parameters and analyze the training datasets used in the machine learning models. Further, the model provides details about the introduction, functional requirements, entropy and uncertainty calculations involved (Confusion Matrix) and regression model used. Additionally, standard semiotics – symbols, icons and indexes are used for user's seamless understanding to use the information system. Below are the details used for developing the web-application.

Client Application Details:

Front end: React JavaScript and TypeScript based Framework.

Languages Used: JS, HTML5 and CSS3.

Server Application Details:

Backend: Node.js Framework for API's development.

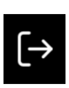
Database: JSON files and Mongo DB.

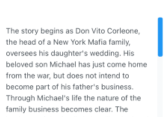
Library Used:

react, react-router-dom, express, cors, mongoose, chart.js, prime-react API's.

2.3 *Semiotics*


Home Button: Icon: This signifies user can navigate to the home page from other page.


Logout Button: Symbol: This signifies user can log off from the application.



Scroll Bar: Symbol: This signifies user can scroll for overflown textual data.

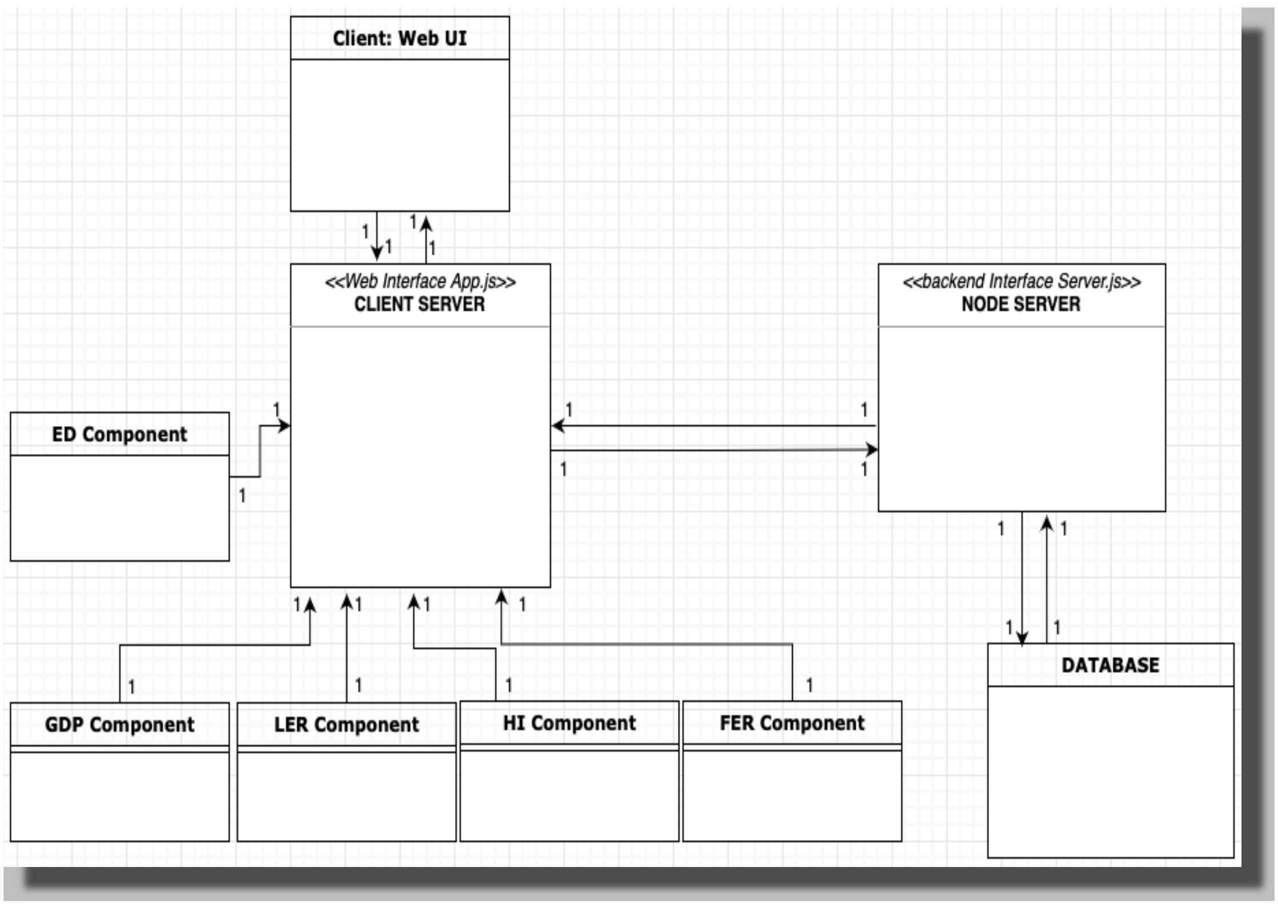
Module Link: Symbol: This signifies user navigate to specific module via clicking it.

2.4 Uncertainty

Given the scope of data procurement from UN, WHO and World Bank by the primary actors for different countries can be a cumbersome process, due to uncertainty in attenuation of data transmission, delays or human error. Hence, uncertainty of data retrieval, processing and model training can high, as improper data can reduce the model accuracy and R2 score future child mortality predictions.

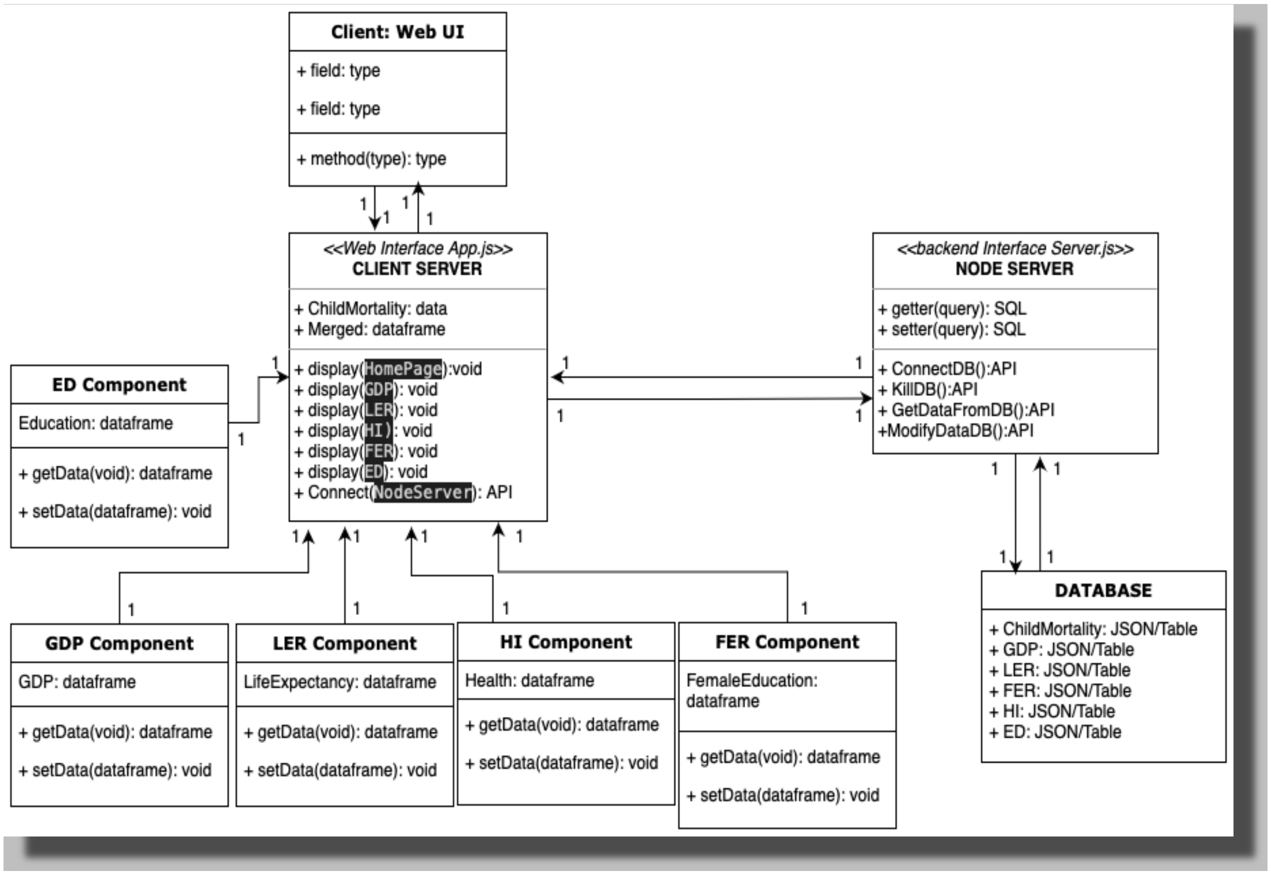
3. THE CONCEPTUAL & LOGICAL MODEL

3.1 Conceptual Model Design



The conceptual model consists of three primary components – Client Web UI, Client Server Interface, and Node Server Interface. These components have bidirectional relationships with each other and mono directional relation with subcomponents.

3.2 Logical Model Design



The Logical model architecture is implemented with below considerations:

- ***GDP Component, LER Component, HI Component, FER Component, ED Component:***
  - Object Oriented Design and Concepts modelling in *JavaScript*.
  - Getter and Setter functions to retrieve and set the data for the client interface.
  - Modular design approach.
- ***Client Server Interface:***
  - Object Oriented Design and Concepts modelling in *JavaScript*.
  - Getter and Setter functions to retrieve the data from subcomponents.
  - Get () and Post () API's request to interact with Node Server Interface.
  - Modular design.
- ***Node Server Interface:***
  - Object Oriented Design and Concepts modelling in *Node.js*.
  - Get () and Post () API's request to interact with database and client server.
  - Modular design.
- ***Database:***
  - Relational database model.
- ***Web UI:***
  - React Framework in JavaScript.
  - Getter functions to connect to Client Sever Interface.

4. DATASETS DESCRIPTION

4.1 HealthCare Expenditure

Child mortality is a very worrying demographic phenomenon especially in developing countries, which has attracted the attention of various researchers and policymakers. Today, combating this issue is considered an important objective, therefore many international organizations, such as the United Nations Children's Fund (UNICEF), the World Bank and the World Health Organization (WHO) have incorporated the objective of reducing child mortality into most of their future programs.[8]

Healthcare financing, whether through private or public means, remains fundamental for the improvement of children's health status all over the world. So, one of the factors having a significant impact on the Global Child Mortality Rate is the Health Expenditures done by the nations across the world.

The role of health economics today is crucial because of growing international awareness of the close relationship between economic development and health. Furthermore, as health in childhood is one of the key predictors of health and productivity in later life, child mortality is an important indicator of socioeconomic development. In the regions, where health infrastructure is largely underdeveloped, increasing health expenditure will contribute to progress towards reducing the child mortality rates. Therefore, in order to achieve the same, the governments in those regions need to increase amounts allocated to health care service delivery. [9]

Globally, the health spending is highly unequal. Hence, the main objective of this effort is to assess the impact of health expenditure on child mortality, as measured by under-five child mortality rates in countries with 50 highest child mortality rates for the period of 2000 to 2015. In addition, we also aim to predict the same using supervised machine learning models.[9]

The data source for this study is the World Health Organization's Global Health Expenditure Database (GHED). This data in the GHED is collected by the WHO and is availed publicly by the World Bank Group according to the open data standards and licenses datasets under the Creative Commons Attribution 4.0 International license (CC-BY 4.0). They are labeled accordingly, and when they are accessed by users, users agree to comply with all of the terms of the respective licenses. [16] GHED is the source of the health expenditure data republished by the World Bank and the WHO Global Health Observatory. The World Bank Group also quotes that in some cases it is not possible to make data available, either because the data are too sensitive, or have been lost or damaged. However, users may still benefit from the available metadata for these datasets.[9]

This data helps ensure health services are available and affordable when people need them. In particular, the data published here contribute to a better understanding of:

- How much do different countries spend on health?
- What are the financing arrangements to pay for health?

WHO works collaboratively with Member States and updates the database annually using available data such as health accounts studies and government expenditure records and where necessary, modifications and estimates are made to ensure the comprehensiveness and consistency of the data across countries and years.[10]

This dataset comprises of Health Expenditure per Capita data for around 190 countries in the world for the time period of 2000 to 2018. This data is collected in local currencies for different nations and then is converted to US$ as per applicable and acceptable norms and conditions. The amount estimates of the current health expenditures include healthcare goods and services consumed during each year by both private and public organizations.

	Country Name	Country Code	Indicator Name	Indicator Code	2000	2001	2002	2003	2004	2005	...	2011
0	Aruba	ABW	Current health expenditure per capita (current...	SH.XPD.CHEX.PC.CD	NaN	NaN	NaN	NaN	NaN	NaN	...	NaN
1	Afghanistan	AFG	Current health expenditure per capita (current...	SH.XPD.CHEX.PC.CD	15.803164	15.803164	15.803164	17.035744	20.412764	23.890501	...	50.853474
2	Angola	AGO	Current health expenditure per capita (current...	SH.XPD.CHEX.PC.CD	12.998967	28.918121	29.049364	34.875187	49.810741	54.260777	...	122.107231

Screenshot of the Health Expenditure Dataset

4.2 Education Index

Education Index is a component of Human Development Index which is published each year by the United Nations. This education index is combined with GDP Index and Life Expectancy Index to measure the GDP per capita and life expectancy.
Education Index is measured by combining the average of average adult years of schooling and expected years of schooling for children. Both factors have equal 50% share each in it.

Education Index (EI):

EI = (MYSI + EYSI) / 2

- Mean Years of Schooling Index (MYSI): MYS/15
  Mean years of schooling (MYS) is a calculation of the average number of years of education received by people ages 25 and older in their lifetime based on education attainment levels of the population converted into years of schooling based on theoretical duration of each level of education attended.
  Fifteen is the projected maximum of this indicator for 2025.

- Expected Years of Schooling Index (EYSI): EYS/18
  Expected Years of Schooling is a calculation of the number of years a child is expected to attend school, or university, including the years spent on repetition. It is the sum of the age-specific enrollment ratios for primary, secondary, post-secondary non-tertiary and tertiary education and is calculated assuming the prevailing patterns of age-specific enrollment rates were to stay the same throughout the child's life.
  Expected years of schooling are capped at 18 years.
  Eighteen is equivalent to achieving a master's degree in most countries. [11]

Globally, the education index is not uniform. Hence, the main objective of this effort is to assess the impact of education index on child mortality, as measured by under-five child mortality rates in countries with 50 highest child mortality rates for the period of 2000 to 2015. In addition, we also aim to predict the same using supervised machine learning models. Thus, investments in education may be a cost-effective means of achieving good health.

The data source for this study is the United Nations Development Programme's Education Index. The Education Index is calculated by UNESCO Institute for Statistics (2020).

This data will help in increasing the access of education to the regions where education index is low.

The data published here contribute to a better understanding of:

- How Education Index varies around the world?
- What are the countries/ regions that have low Education Index and require improvement?

As the time period information is on different axis in both the datasets i.e., on available in rows in the Education Index dataset and available in a column in the Child Mortality dataset, the Education Index dataset is melted & transformed using suitable python libraries in order to make both the datasets similar and compatible. [12]

Multiple data pre-processing techniques are applied for exploratory data analysis and then this dataset is merged with the Child Mortality data to find its one-to-one correlation with it by applying suitable machine learning models to predict the child mortality rate for the all the countries for future years, given the estimated values of the Education Index for that time period.

	Country	Year	Deaths	Education Index
0	Afghanistan	2000	126578.92	0.235
1	Afghanistan	2001	131386.92	0.247
2	Afghanistan	2002	124658.00	0.259
3	Afghanistan	2003	117839.10	0.271
4	Afghanistan	2004	117196.50	0.302

Screenshot of the Dataset Formed After Fusing EI & CM Datasets

4.3 Female Mean Years of Schooling

Mean years of schooling is a calculation of the average number of years of education received by female in their lifetime based on education attainment levels of the population converted into years of schooling based on theoretical duration of each level of education attended. ***SDG 4*** is 'to ensure inclusive and equitable quality education and promote lifelong learning opportunities for all'.

Furthermore, children education is one of the key factors for socio-economic development of a nation, therefore improving the quality of education and providing the access of education to every female child will help in reducing the child mortality rates. Therefore, regions where access to education is very low or poor, improving it will help in addressing the issue of child mortality.

Globally, the education quality and years of schooling varies very much especially due to gender discrimination. Hence, the main objective of this effort is to analyze the impact of female education on child mortality, as measured by under-five child mortality rates in countries with 50 highest child mortality rates for the period of 2000 to 2015. In addition, we also aim to predict the same using supervised machine learning models. [12]

The data source for this study is the United Nations Development Programme's Expected Years of Schooling data. This data is collected by the UNESCO Institute of Statistics, ICF Macro Demographic and Health surveys and UNICEF multiple indicator cluster surveys (2019b).

This data will help in increasing and improving the access of education to female child. In particular, the data will give better understanding of:

- Which countries have highest and lowest years of schooling?
- What is the mean year of schooling of female child?

As the time period information is on different axis in both the datasets i.e., it is available in rows in the Mean Years of Schooling dataset and available in a column in the Child Mortality dataset, the Health Expenditure dataset is melted & transformed using suitable python libraries in order to make both the datasets similar and compatible.
Various data pre-processing techniques are applied for exploratory data analysis and then this dataset is merged with the Child Mortality data to find its one-to-one correlation with it by applying suitable machine learning models to predict the child mortality rate for the all the countries for future years, given the mean years of schooling (female).

	Country	Year	Deaths	Mean Years of Schooling
2	Afghanistan	2010	102919.0	7.2
3	Afghanistan	2011	99331.0	7.4
4	Afghanistan	2012	95248.0	7.5
5	Afghanistan	2013	89519.4	7.6
6	Afghanistan	2014	85306.3	7.7
7	Afghanistan	2015	82957.0	7.8

Screenshot of the dataset formed after merging the MYS and CM datasets

4.4 Life Expectancy Data

Life Expectancy and Child Mortality are very closely related. Child mortality rate plays a very important role in determining the overall life expectancy of a country, this is because a large number of child fatality cancels out the old age people and thus brings down the average lifespan.

Over the last 50 years, the general trend over the world has been to move towards better health and living. All countries, irrespective of the economic condition, always make significant efforts to improve the health of their citizens. One of the major indicators of a country's health well-being is the average Life Expectancy at birth as it reflects the overall mortality index of a population across all the age groups. [14]

The dataset used for our analysis in this project was sourced from The World Bank database. [15] The dataset contained information of most of over 100 countries with the value of Life Expectancy in each for the period of 1950 – 2019. For our analysis, we specifically focused on the top 50 countries with the worst Child Mortality Index and considered the data from the period 2000 – 2015.

Entity	Code	Year	Life expectancy
Afghanistan	AFG	1950	27.638
Afghanistan	AFG	1951	27.878
Afghanistan	AFG	1952	28.361
Afghanistan	AFG	1953	28.852
Afghanistan	AFG	1954	29.35
Afghanistan	AFG	1955	29.854
Afghanistan	AFG	1956	30.365
Afghanistan	AFG	1957	30.882
Afghanistan	AFG	1958	31.403
Afghanistan	AFG	1959	31.925
Afghanistan	AFG	1960	32.446
Afghanistan	AFG	1961	32.962

Screenshot of the Life Expectancy Data Set

For the analysis on this dataset, the data was first pruned to remove any unwanted data i.e., data from the year before the year 2000 and after the year 2015. After the pruning, the data was cleared to remove the rows of the countries that were not part of the considered list. Various data pre-processing techniques are applied for exploratory data analysis and then this dataset is merged with the Child Mortality data to find its one-to-one correlation with it by applying suitable machine learning models to predict the child mortality rate for the all the countries for future years, given the life expectancy at birth.

	Location	Year	Deaths	Life expectancy
0	Afghanistan	2000	126578.92	55.841
1	Afghanistan	2001	131386.92	56.308
2	Afghanistan	2002	124658.00	56.784
3	Afghanistan	2003	117839.10	57.271
4	Afghanistan	2004	117196.50	57.772
5	Afghanistan	2005	117360.00	58.290
6	Afghanistan	2006	120168.00	58.826
7	Afghanistan	2007	112389.00	59.375
8	Afghanistan	2008	109295.00	59.930
9	Afghanistan	2009	108029.00	60.484

Screenshot of the dataset formed after merging the Life Expectancy and Child Mortality Data

4. THE PROTOTYPE IMPLEMENTATION

Functional Requirements:

- Data retrieval from UN, World Bank, and WHO databases for Gross Domestic Product, Life Expectancy Rate, HealthCare Expenditure Index, Education Index and Average Female Education Years.
- Data processing and exploratory analysis on Child Mortality and Target datasets.
- Regression Model Application on merged dataset for predictive analysis country-wise.

Model Used:

```
# import the regressor
from sklearn.tree import DecisionTreeRegressor
from sklearn.ensemble import AdaBoostRegressor

# create a regressor object
regressor1 = DecisionTreeRegressor(max_depth = 3, criterion = "friedman_mse", random_state = 0)
regressor2 = AdaBoostRegressor(DecisionTreeRegressor(max_depth=3, criterion = "friedman_mse"),
                          n_estimators=72, random_state=0)
```

```
Mean absolute error = 4497.7
Mean squared error = 23601575.42
Median absolute error = 4411.3
Explain variance score = 0.98
R2 score = 0.88
```

- Data Visualizations for the analysis on target datasets and predictions.
- Report delivery to primary receivers such as UN and government health body for respective countries.

Non-Functional Requirements:

- Data and predictions report publishing in the MISH server for future references.
- Extending implementation to existing Information Systems.
- Extending the scope to countries.
- Implementing different visualization charts for the analysis.

Evaluations:

The Information System has been weighted against below technical/use case factors and following actions were undertaken:

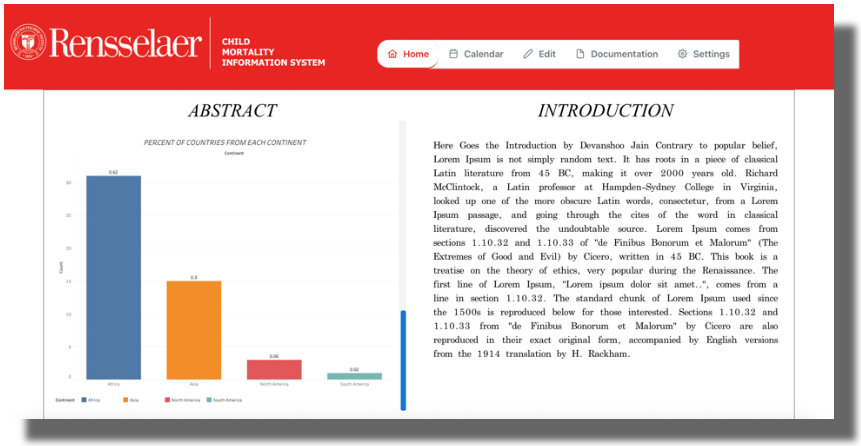
<u>Technical:</u>

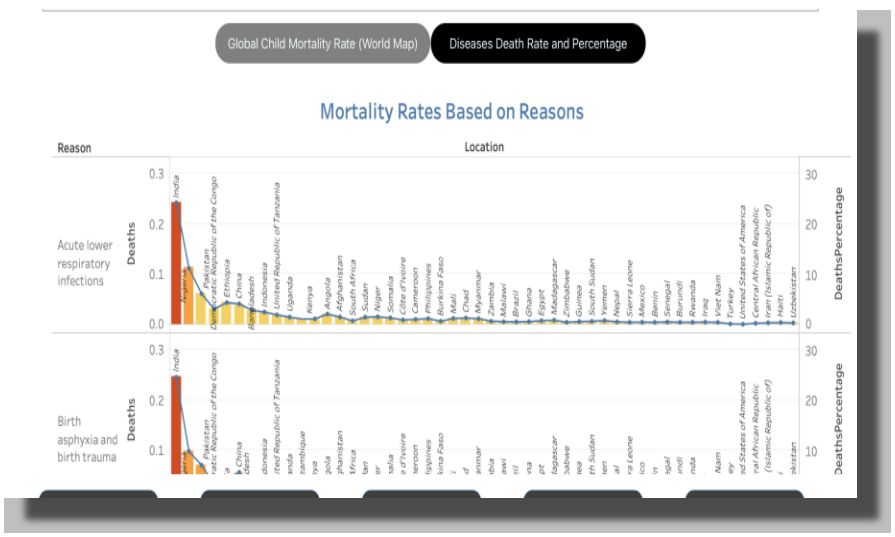
- Data Storage: Databases used were decided on both RDBMS and No SQL databases. As data used don’t consume high volume storage, we have used mongoose db for the requirement.
- Scalability: The information system is scalable to fit against wide and small screen aspect ratios.
- Code Maintainability: As the model has been built with React Framework, addition of new components to the existing feature is quite easy

<u>Use Case:</u>

- New Target variables: As CMIS has dependency on factors such as Healthcare Expenditure Index and Average Female Years of Education, both modules have been adapted to the Information System.

- Data Portability: In order to use the previous years' estimation and analysis MISH cloud server was established.





Screenshots from the client web application

5. CONCLUSION AND FUTURE WORK

- As part of this use case, we developed a Health Information System showcasing the relationshipbetween various factors such as Healthcare Expenditure, Life Expectancy, Education index etc.on Child Mortality for 50 countries over the period of year 2000-2015.

- Various visualizations were plotted that revealed the relationship of these factors with the ChildMortality for each country. The countries India, Nigeria and Pakistan are the top three countries, in respective order, with the highest Child Mortality Ratio considering various reasons for infant fatality.

- The major causes of deaths at the country level were identified to be avoidable diseasessuch as Sepsis, Tetanus, Acute respiratory diseases and Prematurity.

- Using various Machine Learning techniques and algorithms, a predictive model could also estimate the child mortality index of a country given proper data for the aforementioned factors.

- The system could be extended to include more than the current 50 countries to analyze and predict the Child Mortality Data.

- For now, we have focused on one age group (0-4 years) for different nations. The scope of the system could be broadened to include the mortality rate of all age groups.

- The datasets used in this information system are static and were downloaded from World Bank, UN, and WHO etc. This can be extended by dynamically fetching live data using various APIs provided by these sources to create real-time dashboards.

- Furthermore, we can include more parameters like ‘Access to safe drinking water’, ‘Child Immunization Rate’, ‘Environmental factors’ etc. to our existing models to improve predictions of the child mortality index.

- In conclusion, the system provides useful information and insights for various governmental bodies and health organizations to draw conclusions and introspect the reasons for their growing Child Mortality Index and take necessary actions.

- Since child mortality is a growing concern amongst developing and third world nations, our system aims to catalyze the process towards reduced child mortality and a healthier life for all.

6. REFERENCES

[1] https://en.wikipedia.org/wiki/Population_growth

[2] https://www.worlddata.info/average-age.php

[3] https://sdgs.un.org/goals/goal3

[4] https://www.investopedia.com/terms/g/gdp.asp

[5] https://data.worldbank.org/indicator/SH.XPD.CHEX.GD.ZS

[6] https://reactjs.org

[7] https://www.un.org/en/development/desa/population/publications/mortality/world-mortality-cdrom-2019.asp

[8] https://www.nih.gov

[9] https://data.worldbank.org/indicator/SP.DYN.LE00.IN

[10] https://data.worldbank.org/indicator/NY.GDP.MKTP.CD

[11] https://data.oecd.org/gdp/gross-domestic-product-gdp.htm

[12] http://hdr.undp.org/en/data

[14] https://ourworldindata.org/grapher/life-expectancy-vs-health-expenditure

[15] https://link.springer.com/article/10.1007/s10754-019-09272-z

[16] https://apps.who.int/nha/database