

Bad effects of urbanization and lifestyles, Population Health improvements using Analytics.

Varun VS¹, Peter D Augustine²

¹Student, Department of Computer Science, Christ University, Bengaluru, Karnataka, India.

²Associate Professor, Department of Computer Science, Christ University, Bengaluru, Karnataka, India.

-----***-----
Abstract - Healthcare of any country or in general a population is very important in evaluation or measurement of health. If there is enough data or information collected on individual patients, it would provide us a way in either **evaluation** or measurement of health.

There is also a wide range of new diseases which are spreading due to sudden lifestyle changes of a population in any country. We in this research paper try to bring out a prediction of health in the future coming years considering all the lifestyle factors and other aspects of the prevailing occupation of people in a particular population. There is very less physical activity in an urbanized city or a country, There is a requirement now to break the barriers and bring out the bad effects of current population and have a proper measure or medication with the blend of predictive analytics in computer science which can analyze both good and bad effects on a population and either if a population continues with the same lifestyle and predicting bad health in detailed structured way or a population can overcome the barriers and issues of **urbanization** and can implement a measure and improve health much before it gets worsened.

1.INTRODUCTION

According to WHO reports. The urban population in 2014 accounted for 54% of the total global population, up from 34% in 1960, and continues to grow. The urban population growth, in absolute numbers, is concentrated in the less developed regions of the world. It is estimated that by 2017, even in less developed countries, a majority of people will be living in urban areas. The global urban population is expected to grow approximately 1.84% per year between 2015 and 2020, 1.63% per year between 2020 and 2025, and 1.44% per year between 2025 and 2030. [3]

1.1 New cities are spreading infectious diseases

Urban focuses can be impetuses for quick spread of irresistible sicknesses. The premise of extensive populace assembles in a limited region can give the ideal conditions to

various scourges. Global travel has associated the world in the most recent century, and this portability makes a potential risk of many rising sicknesses. Global traveler landings have demonstrated an uncommon development from 25 million in 1950 to 1,087 million in 2013. As indicated by the most recent estimate from the World Tourism Organization, global tourism entries will proceed to increment, and in 2030 the figure is relied upon to be 1.8 billion. With the pace of present day travel, very infectious irresistible sicknesses can be a potential danger in a totally unique setting contrasted with the first flare-up. Urban populace and the thickness of occupants can meet the criteria for another scourge and make a general wellbeing debacle, if not considered important.

1.2 Urban Population looks very Disadvantageous

From a couple of decades, we have seen expanded urbanization as well as very disadvantageous to the urban populace principally youngster's populace. This is fairly striking given the way that in the nineteenth century Britain under-5 mortality seems to have been an essential part of the urban penalty [4].

Given poor people and now and again ghetto conditions in which numerous urban kids are living in low and center salary nations, is there any sign that they are by and by more awful off than youngsters living in non-urban settings? Dependable and comprehensive information on the well-being of urban and rustic populaces is moderately scanty. All things considered, what exists proposes that in many occasions kid wellbeing is better, and mortality lower, in urban than in provincial zones. The nonappearance of an urban punishment in kids wellbeing in low and center salary nations today is the after effect of a blend of variables. To start with, in these nations supreme destitution and ailing health are for the most part rustic marvels. Lack of healthy sustenance is a critical hazard calculate for death from basic adolescence contaminations of diarrhea and respiratory ailment urban conditions will have a tendency to be in an ideal situation. Besides, urban populaces have a tendency to

have higher immunization scope rates than their rustic partners.

Thirdly, scattering and take-up of viable intercessions, for example, oral rehydration treatment are probably going to be more viable in an urban setting. The tyke wellbeing preferred standpoint of urban ranges is not an explanation behind smugness. Kids' living conditions, wellbeing and survival in urban territories in most low and center pay nations are poor contrasted and those in high salary nations.

In addition, as is frequently highlighted, there are huge disparities in kid wellbeing and mortality inside urban communities between poor people and the affluent [5] These are the main cases at kids wellbeing there are numerous different cases as for grown-up wellbeing too.

2. Predictive Analytics

Can Healthcare Industry Really Utilize It Fully and do any prediction on population health?

The popular expression fever around prescient examination will probably proceed to rise and fall. Sadly, deficient with regards to the correct foundation, staffing, and asset to act when something is anticipated with high conviction to happen, we miss the mark concerning the maximum capacity of outfitting noteworthy patterns and examples in patient information. At the end of the day, without the resolution for clinical mediation, any indicator regardless of how good is not completely used. What does this need to do with Health Catalyst and enterprise data appropriation focus capacity? Prosperity Catalyst not simply has the most ideal development structure and metadata set up, yet we furthermore make novel prescriptive examination. More fundamentally, we join each one purposely to our required needs for instance, cost suitability, clinical traditions or patient outcomes.

Keeping in mind the end goal to be fruitful, we feel that clinical occasion forecast and ensuing intercession ought to be both substance driven and clinician driven. A more specific term is rigorous analytics, which would incorporate evidence, suggestions for each anticipated classification or result.

Why? In prescription, innovation and proof ought to have a similar ultimate objective – to amplify the utility of the chronicled drift for enhancing understanding consideration. Inside Health Catalyst, information displaying and

calculation improvement is performed utilizing industry driving instruments for information mining and regulated machine adapting, for example, Weka, Orange, and R. Progressing endeavors incorporate order models for a summed up indicator of healing facility readmissions, heart disappointment, length of stay and bunching of patient results to chronicled accomplices at time of concede. Above all, we have inward access to a large number of de-distinguished healing facility records in both the inpatient and outpatient settings and grown-up and pediatric populaces. This preparation information is vital to tending to the prescient examination requests of customers and site customization. So when your demand comes whether it includes order or grouping or highlight determination. Health catalyst has the apparatuses and the information and the skill to effectively convey beat performing the prescient investigation.

Collection of Datasets

We in this exploration attempt to draw out an expectation in the matter of how the strength of a populace would be in future with the proceeding with a way of life and different elements. Our datasets will include a great deal of open source clinic information which includes different elements, Alongside this auxiliary datasets, We additionally attempt to gather some essential information from a specific urban populace to draw out an induction and how a specific populace can know about such new issues and be prepared and can likewise focus on wellness angles to maintain a strategic distance from such issues. We shall also make the analysis more patient-centric by focusing on the fitness aspects and providing the patients online proven links of home remedies to general diabetes and hypertension and overall health development through fitness.

Methodology

Data mining can be called as a method to extricate information from data. It is the way toward getting covered up data and afterward transforming it into qualified information by measurable methods, man-made brainpower method. One of its application zones is a therapeutic range to frame choice emotionally supportive networks for analysis just by concocting significant data from given medicinal information [7]. In this concentrate a choice emotionally supportive network for finding of ailment that make utilize of information mining and three distinctive manmade

brainpower classifier calculations to be specific Multilayer Perceptron, Naive Bayes Classifier and J.48. Pima Indian dataset of UCI Machine Learning Repository was utilized. Above mentioned dataset incorporates urinary and blood test consequences for 768 patients. These test outcomes comprise of 8 diverse component vectors. The proposals for future reviews were displayed.

Data Attributes

Dataset is made out of 768 cases,found in Table I. [6]Every person is portrayed in an informational collection of 8 characteristics. These characteristics are numerical qualities.

Table -1: The attributes of dataset

Number	Name	Type
1	Number of times pregnant	Numeric
2	Plasma glucose concentration a 2 hours in an oral glucose tolerance test	Numeric
3	Diastolic blood pressure (mm Hg)	Numeric
4	Triceps skin fold thickness (mm)	Numeric
5	2-Hour Serum insulin (mu U/ml)	Numeric
6	Body mass index (weight in kg/(height in m) ²)	Numeric
7	Diabetes pedigree function	Numeric
8	Age (years)	Numeric

These properties are: Diastolic blood pressure, plasma glucose fixation a 2 hours in an oral glucose resilience test, Diastolic pulse (mm Hg), triceps skin fold thickness(mm), 2-Hour serum insulin (mu U/ml), BMI (weight in kg/(stature in m)²), diabetes ,age (years), Class variable (0 or 1).

Table -2: Statistical data of all features

Attribute	Min value	Max value	Mean Value	Standard Dev.
1	0	17	3.845	3.37
2	0	199	120.895	31.973
3	0	2.42	69.105	19.356
4	0	99	20.536	15.952
5	0	846	79.799	115.244
6	0	67.1	31.993	7.884
7	0.078	122	0.472	0.331
8	21	81	33.241	11.76

a. Classification

Classification is the method of recognizing to which of an arrangement of classes another observation has a place, on the start of a planning set of data containing recognitions whose class interest is known. Cluster analysis strategies are utilized mostly to total items into gatherings as indicated by likeness measures. Many reviews have been made to think about the numerous distinctive techniques for classification.

With this review, three unique classifiers (MLP, J48, and Naive Bayes) are connected with Pima Indians Diabetes Informational collection. We come with an inference with this.

b. Multi-Layer Perceptron

The artificial neural network (ANN) or neural system, in short, is roused by mimicking the capacity of a human mind.

A neural system can be utilized to speak to a nonlinear mapping amongst info and yield vectors. Neural systems are among the prevalent flag preparing innovations. In the development, neural frameworks serve two basic systems as example 1.Classifiers 2.Nonlinear versatile channels.

A general system comprises a layered design, an input layer, at least one shrouded layers and a yield layer

The Multilayer perceptron (MLP) is an imaginary or fake neural system that is utilized widely to illuminate a number of various issues, design acknowledgment, interjection. Every layer is made out of neurons, which are related with each other by weights. In every neuron, a particular numerical capacity called the actuation work acknowledges the contribution from past layers and creates yield for the next layer. Here, the actuation work utilized is the digression sigmoid exchange work.

J48

J48 is an open source Java execution of the C4.5 calculation in WEKA information exploration instrument. C4.5 is a calculation for creating a choice tree created by Ross Quinlan. The choice trees created by C4.5 can be utilized for characterization, so C4.5 is frequently alluded as factual classifier. For initiating characterization runs as Decision Trees from an arrangement of given illustrations C4.5 calculation was presented by Quinlan. C4.5 is a development and refinement of ID3 that represents inaccessible qualities, persistent characteristic esteem ranges, pruning of choice trees, run deduction, etcetera. A set of records is given.

Naive Bayes

The Naive Bayes Classifier procedure is fundamentally suited when the dimensionality of the data sources is high. In spite of its effortlessness, Naive Bayes can regularly beat more refined grouping strategies.

One of the simplest methods for choosing the most plausible speculation given the information that we have that we can use as our earlier learning about the issue. Bayes' Theorem gives a way that we can figure the likelihood of a theory given our earlier information.

3. CONCLUSIONS

Table-3 Classification Results

Number	Classification Types	Accuracy	Time(sec.)
1	MLP	75.130	1.13
2	J48	73.828	0.08
3	Naive Bayes	76.302	0.01

Classification results are in Table III. Regarding outcomes, the chosen methods from specific writing accomplished expectation precision proportions of 75.13%, 73.828%, and 76.302% for MLP, J48, and Naïve Bayes, respectively.

This proves that Naive Bayes has the best execution among the all others on the dataset.

These classifiers are trending in field of medicine and data engineering.

Counter measures to urbanization and Population Health Improvements

A. Coordinate change of urbanization

Urbanization improvement cannot relinquish natural condition, not to the detriment of human well-being, not at the cost of social human progress. On the off chance that urbanization improvement surpasses environmental conveying limit, a progression of biological issues will be formed, so that over-urbanization ought to be limited.

Unexpectedly, moderate urbanization advancement can ensure natural condition. however, it limits industrialization and monetary improvement which will bring down the non-rural level.

In the event that urban development is lingering behind, countless surplus works, society wasteful, low urbanization will be found. A local populace which fits monetary and social advancement level can get great general wellbeing administrations. As well high or too low is undesirable.

B. Economy Re-usage

The attention on costs applies to doctors' wages too. The model doctor's facilities pay specialists a settled compensation, offering no rewards or different motivating forces for inner referrals for tests or methodology—an approach quite examined by social insurance scholars in created nations yet more often than not actualized due to charge for-administration impetuses. The best Indian clinics offer settled costs for a bundle of regular restorative medications as opposed to valuing for individual techniques or tests. The settled value demonstrate disheartens specialists from requesting superfluous methodology, while

conventions guarantee that fundamental strategies aren't skipped.

[8] Data Mining in the Telecommunications Industry, Gary M. Weiss, Fordham University, USA.

C. Effectively fabricating solid city

We feel that alone this city can be called as a sound city where occupants diligent work to enhance the city's common and social condition, to completely investigate the group assets, to commonly bolster each other and expand their potential. Affecting components of Health city incorporate three factors: wellbeing related factors (HRV, for example, wellbeing training, vaccination, ailment control, medicinal services, wellbeing administrations, wellbeing supervision etc.

socio-monetary status variables (SEV, for example, work, wage, populace control, etc.

Living condition variables (LEV, for example, home, water supply, foundation et cetera. At present, solid city's development ought to concentrate on state clean urban communities, solidification, additionally, reinforce infection avoidance and control, enhance wellbeing administration framework and assurance people oriented, population wellbeing focused.

REFERENCES

[1] Student, Department of Computer Science, Christ University, Bengaluru, Karnataka, India. varunvachis6@gmail.com

[2] Associate Professor, Department of Computer Science, Christ University, Bengaluru, Karnataka, India.

[3]http://www.who.int/gho/urban_health/situation_trends/urban_population_growth_text/en/

[4] Peter D Augustine. Article: Leveraging Big Data Analytics and Hadoop in Developing India's Healthcare Services. *International Journal of Computer Applications* 89(16):44-50, March 2014.

[5] Jiawei Han and Micheline Kamber, Data mining, concept and techniques.

[6]<https://archive.ics.uci.edu/ml/datasets/Pima+Indians+Diabetes>

[7] https://en.wikipedia.org/wiki/Data_mining