**A Major Project Report**

**On**

**PREDICTING LIFE EXPECTANCY USING ML**

**Submitted in partial fulfillment of the**

**Requirements for the award of**

**INTERNSHIP CERTIFICATE**

**IN**

**MACHINE LEARNING**

**Submitted By**

**THALLA YUVANA**

**Under the guidance of**

**Mr.CHARAN**

**ACKNOWLEDGEMENT**

I would like to express our sincere thanks to Mr.charan for his motivation and co-operation for the successful completion of the project.

I would like to express our sincere thanks to my ML-6 batch friends for their encouragement to achieve the goals of the project.

-**Yuvana Thalla**

**DECLARATION**

This is to certify that the project work entitles “Predicting life Expectancy” submitted to smartinternz in partial fulfillment of the requirement for the award of Internship certificate is an original work carried out by T.Yuvana under the guidance of Mr.Charan. This matter embodied in this project is a genuine work, done by me.

-**Yuvana Thalla**

Project report

Introduction:

Over view

A typical Regression Machine Learning project leverages historical data to predict insights into the future. The aim is to predict the Life Expectancy rate of a country given various features.

Life expectancy is a statistical measure of the average time a human being is expected to live, Life expectancy depends on various factors: Regional variations, Economic Circumstances, Sex Differences, Mental Illnesses, Physical Illnesses, Education, Year of their birth and other demographic factors. This statement provides a way to predict average life expectancy of people living in a country when various factors such as year, GDP, education, alcohol intake of people in the country, expenditure on healthcare system and some specific disease related deaths that happened in the country are given.

Purpose

Lifeexpectancypredictions have the potential to be beneficial to individuals, health service providers and governments. For instance, they would make people more aware of their general health, and its improvement or deterioration over time. This may motivate them to make healthier lifestyle choices Life expectancy is one of the most important factors in end-of-life decision making. It helps to determine the course of treatment and helps to anticipate the procurement of health care services and facilities, or more broadly: facilitates Advance Care Planning. Advance Care Planning improves the quality of the final phase of life by stimulating doctors to explore the preferences for end-of-life care with their patients, and people close to the patients. This research tests the potential of using machine learning and natural language processing techniques for predicting life expectancy.

Literature survey:

Existing problem

Life expectancy refers to the number of years a person is expected to live based on the statistical average of different factors. Life expectancy varies by birth year, geographical area, by era and many factors.

Life expectancy, estimate of the average number of additional years that a person of a given age can expect to live. The most common measure of life expectancy is life expectancy at birth. Life expectancy is a [hypothetical](https://www.merriam-webster.com/dictionary/hypothetical) measure. It assumes that the age-specific [death](https://www.britannica.com/science/death) rates for the year will apply throughout the lifetime of individuals born in that year. The measure differs considerably by sex, age, race, and geographic location of the person. Therefore, life expectancy is commonly given for specific categories.

There may be so many factors to estimate life of a person more accurately. Such as mental and physical illness, Education, other demographic factors data by the time of death which gives precise Life expectancy values.

Proposed solution

The main aim of project is to create a user interface which can be helpful for predicting life expectancy of particular country. In order to achieve that one we use IBM cloud platform services. IBM cloud platform services provide us to implement machine learning service associated with Watson studio. By using those services we create project in Watson studio. Where we can import our dataset as asset and add Auto AI services to project then we select the predicting column after that we can deploy our model.

Another main step in our project is to create user interface and associate it with deployed ML project which we create in Watson studio. To achieve that, we use Node-RED software in IBM platform. In Node-RED we use form node to create UI, function node for input, IBM node to integrate ml service to UI, response for displaying predicted value.

By considering many factors Data such as Regional variation, Economic circumstances and GDP, Adult morality, Infant deaths, Diseases caused of the country we make this Machine Learning Model which leverages these historic data in predicting the insights into the future.

Theoretical Analysis:

Block diagram

Predicting life expectancy

Build node red flow to integrant Auto- AI model

Deploy and test your model

Add the data set to the Auto- AI experiment and run

Create a new project in Watson studio

Collect the dataset

Hardware/Software designing

* processor
* IBM cloud account, Dataset, IBM Watson service, Watson studio, Machine Learning service, Auto AI Experiment and Node-RED software etc.

Experimental investigations:

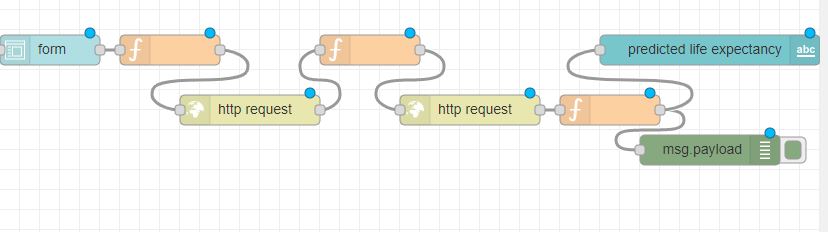


Fig 1: node-red flow

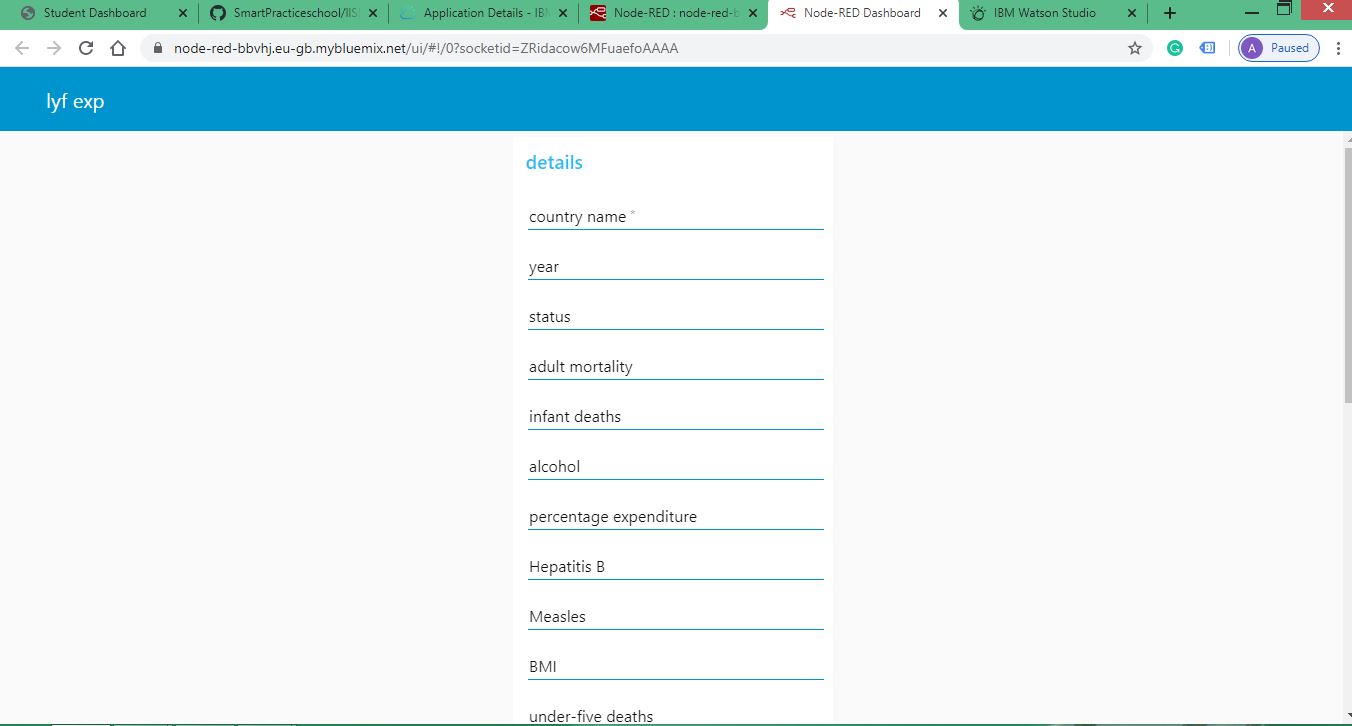


Fig 2: user interface

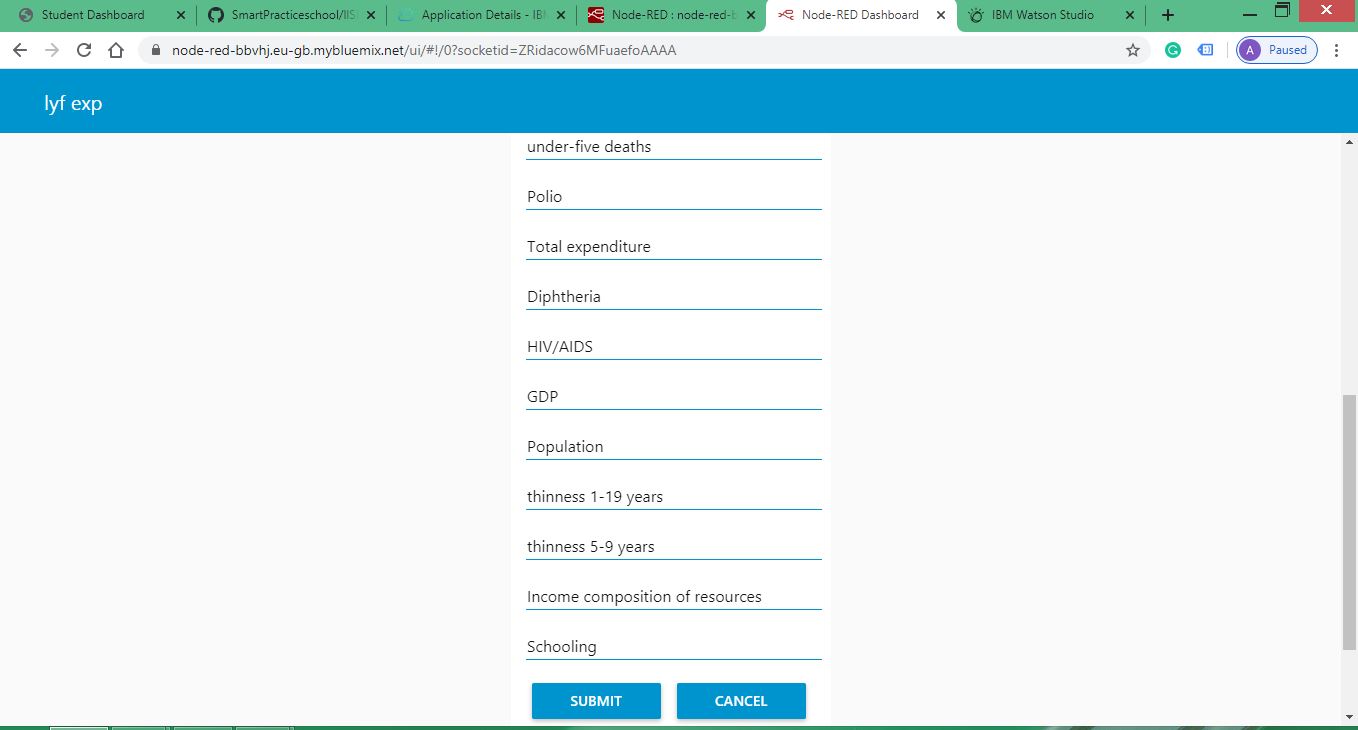


Fig 3: user interface

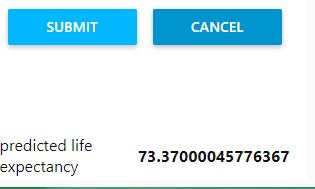


Fig 4: output

Flow chart:

Enter values

Predicting life expectancy of a country

Display predicted value

Result:

Life expectancy is predicted by using with the help of various factors such as Regional variations, Economic Circumstances, Sex Differences, Mental Illnesses, Physical Illnesses, Education, Year of their birth and other demographic factors.

This project will predict average life expectancy of people living in a country when various factors such as year, Population, infant deaths, GDP, education, alcohol intake of people in the country, expenditure on healthcare system and Income composition of resources and other factors are given.

Advantages and Disadvantages:

Advantages

Accuracy. Life expectancy using life insurance underwriting and life settlement methods is very accurate, which can affect millions of dollars in judgments or awards. This addresses the inadequacy of life expectancies that are simply taken from a life table, estimated by a doctor, actuary or statistician, or have been excluded. In many cases, such life expectancies are too short, too long and/or not credible.

 Provides a complete insight to the judge or jury of the individual's medical conditions and/or personal history of high-risk behaviors. A life expectancy using life insurance underwriting and life settlement methods provides a comprehensive profile of the individual's medical conditions as well as any risky behaviors. The judge or jury will be given a comprehensive view of the individual's personal health and lifestyle, which often decides the case.

Disadvantages

The main disadvantage is that NO ONE can predict the future. No one knows when someone will die, who will get cancer or not, who will recover and who won't. A person who appears to have been sickly for many years can surprise everyone by outliving all of his peers. Likewise, a person known to be robustly healthy and in shape can succumb immediately to a stroke, heart attack, or other unexpected calamity.

If you use life expectancy, it doesn't mean that you or anyone else will live as long as the chart says they will. Even if their health, sex, age, race and medical condition are all the same as the charts, doesn't mean that they will live as long as the chart says. That is because life expectancy is an average life span of thousands of people. That means that some of them died much earlier, and some died much later than the expected average life span.

Applications:

Along with existing heath applications such as fitness tracking, chronic disease monitoring and real-time patient monitoring, the PLE application can be useful for users to improve their lifestyle and exercise by planning goals on a short and long-term basis.

\*The development of wearable devices is evolving rapidly to capture data and for use in applications. By wearing fitness tracking or monitoring devices, those attributes can be automatically updated and sent to the cloud servers, which process and update the health index defined by the user’s physician to assess the user’s health life, which eventually leads to updating the user’s PLE.

\*One of the potential stakeholders of this solution may be health insurance companies, who can provide personalized insurance products by monitoring their customers’ health index. They may provide customers with incentives to improve their health condition and lifestyle, for example by revising their premiums in response to improvements in their customers’ health index obtained and processed by the cloud computing program.

Conclusion:

This concluding part briefly summarizes the context in which the study of regional mortality differences is embedded. It highlights some important scientific contributions of this study, reflects on its shortcomings, and offers suggestions for future research. In addition, some potential mortality scenarios are outlined.

Studying the regional forerunners of low mortality illustrates the potential for global pathways to increased longevity. The potential for mortality reductions is, in some respects, obvious. The study showed that the regions with the highest life expectancy are not necessarily the forerunners in the reduction of unnecessary deaths.

The principal focus of this study was on identifying regional mortality trends over time, as well as the reasons behind these differentials.

Future scope:

Future studies should attempt to make the effects of lifestyle and health-care factors on regional mortality differences more explicit, as they are more proximate determinants than socioeconomic factors. From a theoretical point of view, more meaningful indicators reflecting the quality and accessibility of health care at the contextual levels should be developed in order to identify deficiencies in the health care system. At the individual level, the fact that the privately insured live longer than those in compulsory health insurance, independent of their socioeconomic status, is striking. This raises the question of whether prevention and medical care are better for those who are privately insured or whether there are distinct selection effects into private health.

Source code:

[{"id":"b4c1e086.6150a","type":"tab","label":"life expectancy","disabled":false,"info":""},{"id":"6c073639.6d65b8","type":"function","z":"b4c1e086.6150a","name":"","func":"global.set(\"country\",msg.payload.country)\nglobal.set(\"yr\",msg.payload.yr)\nglobal.set(\"sts\",msg.payload.sts)\nglobal.set(\"am\",msg.payload.am)\nglobal.set(\"id\",msg.payload.id)\nglobal.set(\"alc\",msg.payload.alc)\nglobal.set(\"pe\",msg.payload.pe)\nglobal.set(\"hep\",msg.payload.hep)\nglobal.set(\"mes\",msg.payload.mes)\nglobal.set(\"bmi\",msg.payload.bmi)\nglobal.set(\"ufd\",msg.payload.ufd)\nglobal.set(\"pol\",msg.payload.pol)\nglobal.set(\"te\",msg.payload.te)\nglobal.set(\"dph\",msg.payload.dph)\nglobal.set(\"ha\",msg.payload.ha)\nglobal.set(\"gdp\",msg.payload.gdp)\nglobal.set(\"pop\",msg.payload.pop)\nglobal.set(\"thin\",msg.payload.thin)\nglobal.set(\"thinn\",msg.payload.thinn)\nglobal.set(\"ic\",msg.payload.ic)\nglobal.set(\"scl\",msg.payload.scl)\nvar apikey=\"TccYWxhIh5-X8Xi42csCUzY9hT79Vs6aMN2-uViMoVPm\";\nmsg.headers={\"content-type\":\"application/x-www-form-urlencoded\"}\nmsg.payload={\"grant\_type\":\"urn:ibm:params:oauth:grant-type:apikey\",\"apikey\":apikey}\nreturn msg;","outputs":1,"noerr":0,"x":190,"y":260,"wires":[["968e529c.fb426"]]},{"id":"968e529c.fb426","type":"http request","z":"b4c1e086.6150a","name":"","method":"POST","ret":"obj","paytoqs":false,"url":"https://iam.cloud.ibm.com/identity/token","tls":"","persist":false,"proxy":"","authType":"","x":270,"y":320,"wires":[["8acf15f7.5f8948"]]},{"id":"fe313dce.6a8cf","type":"debug","z":"b4c1e086.6150a","name":"","active":true,"tosidebar":true,"console":false,"tostatus":false,"complete":"payload","targetType":"msg","x":730,"y":360,"wires":[]},{"id":"8acf15f7.5f8948","type":"function","z":"b4c1e086.6150a","name":"","func":"var country = global.get('country')\nvar yr = global.get('yr')\nvar sts = global.get('sts')\nvar am = global.get('am')\nvar id = global.get('id')\nvar alc = global.get('alc')\nvar pe = global.get('pe')\nvar hep = global.get('hep')\nvar mes = global.get('mes')\nvar bmi = global.get('bmi')\nvar ufd = global.get('ufd')\nvar pol = global.get('pol')\nvar te = global.get('te')\nvar dph = global.get('dph')\nvar ha = global.get('ha')\nvar gdp = global.get('gdp')\nvar pop = global.get('pop')\nvar thin = global.get('thin')\nvar thinn = global.get('thinn')\nvar ic = global.get('ic')\nvar scl = global.get('scl')\nvar token=msg.payload.access\_token\nvar instance\_id=\"3ce2a322-5d19-4499-8624-9ed72aa7cdcf\"\nmsg.headers={'Content-Type': 'application/json',\"Authorization\":\"Bearer \"+token,\"ML-Instance-ID\":instance\_id}\nmsg.payload={\"input\_data\": [{\"fields\": [\"Country\", \"Year\", \"Status\", \"Adult Mortality\", \"infant deaths\", \"Alcohol\", \"percentage expenditure\", \"Hepatitis B\", \"Measles \", \" BMI \", \"under-five deaths \", \"Polio\", \"Total expenditure\", \"Diphtheria \", \" HIV/AIDS\", \"GDP\", \"Population\", \" thinness 1-19 years\", \" thinness 5-9 years\", \"Income composition of resources\", \"Schooling\"], \"values\":[[country,yr,sts,am,id,alc,pe,hep,mes,bmi,ufd,pol,te,dph,ha,gdp,pop,thin,thinn,ic,scl]] }]}\nreturn msg;","outputs":1,"noerr":0,"x":390,"y":260,"wires":[["2ad2dabd.2ebed6"]]},{"id":"2ad2dabd.2ebed6","type":"http request","z":"b4c1e086.6150a","name":"","method":"POST","ret":"obj","paytoqs":false,"url":"https://eu-gb.ml.cloud.ibm.com/v4/deployments/e8c7d47a-6874-443f-aa2c-c610db4b528d/predictions","tls":"","persist":false,"proxy":"","authType":"","x":490,"y":320,"wires":[["d8740f09.65c17"]]},{"id":"11cb343.013b5cc","type":"ui\_form","z":"b4c1e086.6150a","name":"","label":"","group":"695ac23d.e3b9dc","order":1,"width":0,"height":0,"options":[{"label":"country name","value":"country","type":"text","required":true,"rows":null},{"label":"year","value":"yr","type":"number","required":false,"rows":null},{"label":"status","value":"sts","type":"text","required":false,"rows":null},{"label":"adult mortality","value":"am","type":"number","required":false,"rows":null},{"label":"infant deaths","value":"id","type":"number","required":false,"rows":null},{"label":"alcohol","value":"alc","type":"number","required":false,"rows":null},{"label":"percentage expenditure","value":"pe","type":"number","required":false,"rows":null},{"label":"Hepatitis B","value":"hep","type":"number","required":false,"rows":null},{"label":"Measles","value":"mes","type":"number","required":false,"rows":null},{"label":"BMI","value":"bmi","type":"number","required":false,"rows":null},{"label":"under-five deaths","value":"ufd","type":"number","required":false,"rows":null},{"label":"Polio","value":"pol","type":"number","required":false,"rows":null},{"label":"Total expenditure","value":"te","type":"number","required":false,"rows":null},{"label":"Diphtheria","value":"dph","type":"number","required":false,"rows":null},{"label":"HIV/AIDS","value":"ha","type":"number","required":false,"rows":null},{"label":"GDP","value":"gdp","type":"number","required":false,"rows":null},{"label":"Population","value":"pop","type":"number","required":false,"rows":null},{"label":"thinness 1-19 years","value":"thin","type":"number","required":false,"rows":null},{"label":"thinness 5-9 years","value":"thinn","type":"number","required":false,"rows":null},{"label":"Income composition of resources","value":"ic","type":"number","required":false,"rows":null},{"label":"Schooling","value":"scl","type":"number","required":false,"rows":null}],"formValue":{"country":"","yr":"","sts":"","am":"","id":"","alc":"","pe":"","hep":"","mes":"","bmi":"","ufd":"","pol":"","te":"","dph":"","ha":"","gdp":"","pop":"","thin":"","thinn":"","ic":"","scl":""},"payload":"","submit":"submit","cancel":"cancel","topic":"","x":70,"y":260,"wires":[["6c073639.6d65b8"]]},{"id":"5df788ed.15f258","type":"ui\_text","z":"b4c1e086.6150a","group":"695ac23d.e3b9dc","order":2,"width":0,"height":0,"name":"","label":"predicted life expectancy","format":"{{msg.payload}}","layout":"row-spread","x":730,"y":260,"wires":[]},{"id":"d8740f09.65c17","type":"function","z":"b4c1e086.6150a","name":"","func":"msg.payload=msg.payload.predictions[0].values[0][0]\nreturn msg;","outputs":1,"noerr":0,"x":630,"y":320,"wires":[["fe313dce.6a8cf","5df788ed.15f258"]]},{"id":"695ac23d.e3b9dc","type":"ui\_group","z":"","name":"details","tab":"9fa339fb.d0fc08","order":3,"disp":true,"width":"6","collapse":false},{"id":"9fa339fb.d0fc08","type":"ui\_tab","z":"","name":"lyf exp","icon":"le","order":1,"disabled":false,"hidden":false}]