**Diabetes** **Detector**

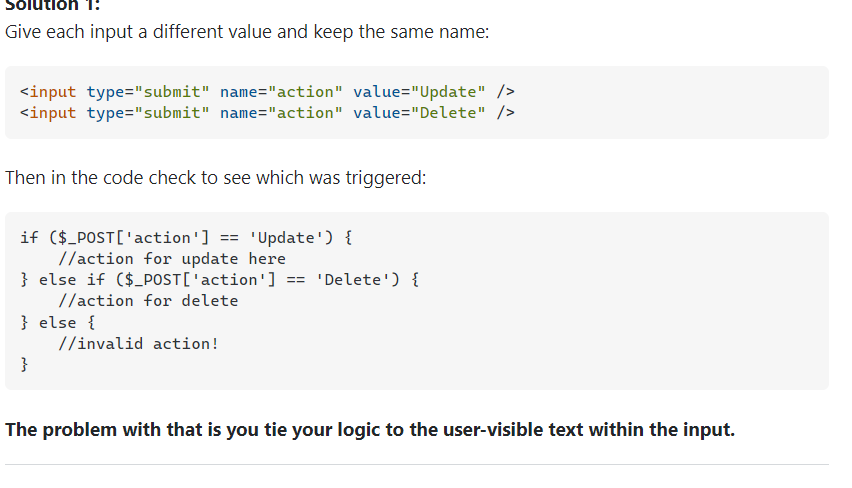
* First create a website with a form that can take input values
* Then train a model using the dataset provided
* Then predict the values using the model that you trained
* And then renter the data predicted in the entered the data to predict the model
* Make the site friendly and creative

Progress so far:

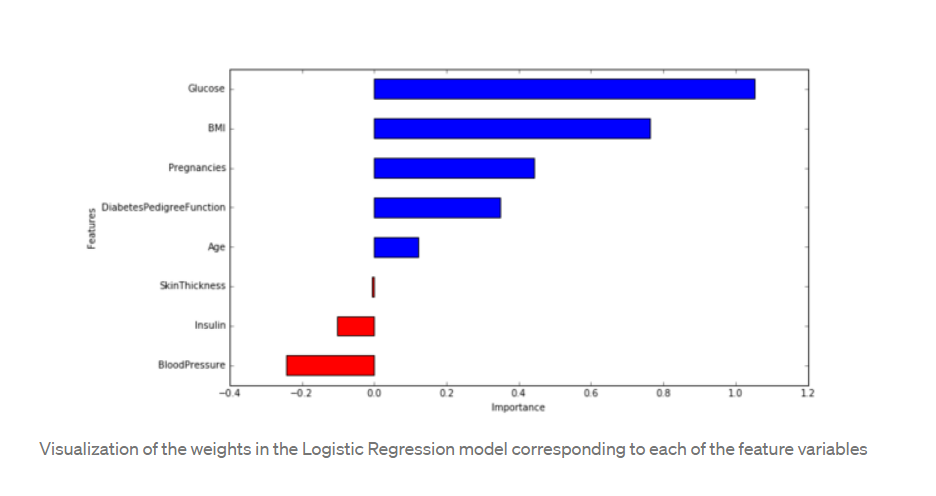
* First page done
* Anchor link to form
* Data on first page

Error so far:

* Website mein reurn karte sami error aara to return render template mein hi karna pagega



Effect of Parameters on ML model



1. Glucose level, BMI, pregnancies and diabetes pedigree function have significant influence on the model, specially glucose level and BMI. It is good to see our machine learning model match what we have been hearing from doctors our entire lives!
2. Blood pressure has a negative influence on the prediction, i.e. higher blood pressure is correlated with a person not being diabetic. (also, note that blood pressure is more important as a feature than age, because the magnitude is higher for blood pressure).
3. Although age was more correlated than BMI to the output variables (as we saw during data exploration), the model relies more on BMI. This can happen for several reasons, including the fact that the correlation captured by age is also captured by some other variable, whereas the information captured by BMI is not captured by other variables.

So I can say that

1. Glucose , Bmi , pregnancies,dpf and age are directly proportional to chances of diabetes.
2. Skin thickness , insulin , Bp are inversely proportional to chances of diabetes.

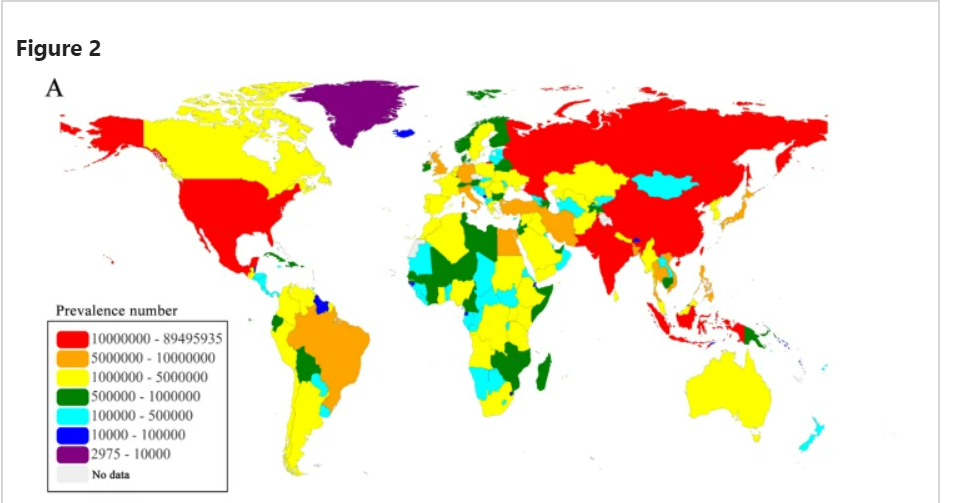
What to do next

1. You could also improve the data cleaning, by replacing 0 values with the mean value.
2. Add more routes for getting the location.
3. Make different models for male and females and add regions probability to them.
4. Change photos and add new photos

Trends of diabetes region wise

1. Age-standardized rates of incidence, prevalence, death, DALYs were expressed as number per 100,000 population.
2. Means the survey were in a 100,000 population
3. The five highest prevalence were observed in China (89.5 million), India (67.8 million), United States (30.7 million), Indonesia (21.0 million), and Mexico (13.1 million).

Prevalence Number



Distribution

1. Red :
   1. Russia , India, Pakistan, China, U.S.A ,Indonesia
2. Yellow:
   1. Canada, Mexico, Columbia , Venezuela, Chile , Peru, Argentina, **French Guiana**, Africa( South africa, angola, Dom. Repu. Of congo, Uganda, Tanzania, kenya, sudan, Ethiopia, ghana, Nigeria, Algeria, Morocco, yemen, suadi arabia,),iraq,Syria, Europe(spain , Portugal , France, Belgium , Netherlands, Greece, Romania,Ukraine,Sweden,),Asia(kazakhastan,Uzbekistan,afghanistan,Nepal,sri lanka,Malaysia,Burma,Australia,n.korea,s.korea,)
3. Orange:
   1. Brazil, egypt, Italy, Germany, u.k, Iran, turkey, Japan, Bangladesh,Vietnam, phillipenes,
4. Green:
   1. Carribian, Ecuador, Bolivia, cote d’lvoire, b.faso, mali, **Nigeria** , Libya, Senegal, gambia, Zambia, zimbawe, Malawi , mozambigue, madagasker, Somalia, Jordan, U.A.E , Az. , Bulgaria, Belarus, Norway, Ireland, Tajakistan, **Norway**, Combodia, Papua new guinia
5. Purple:
   1. Greenland
6. Dark Blue:
   1. Iceland, Guyana**, French Guiana**, suiriname, guinea-bissau , equaltorial guinea, dijibouti, eswatini, Bhutan
7. Light Blue:
   1. **Central America**, Uruguay, Paraguay, Namibia, bots(Africa), rawanda,Burundi,eritria,oman,Qatar,Kuwait,Liberia,**meurit**, sierra leone, guinea, **ghana**, gabon,congo,c.a.r, chad, south sudan, Lesotho, Estonia, Latvia,Lithuania,Georgia, turkm, kyrgystan, Mongolia, laos, fiji , new Zealand

What I Think:

1. Because the survey was in u.s then set it as benchmark and then compare others problem solved