Annual survey data of 400k adults related to their health status:

The original dataset of nearly 300 variables was reduced to just about

18 variables. In addition to classical EDA, this dataset can be used to apply a range of machine learning methods, most notably classifier models. You should treat the variable "HeartDisease" as a binary ("Yes" - respondent had heart disease; "No"

- respondent had no heart disease). But note that classes are not balanced, So the classic model application approach is not advisable.

Fixing the weights/undersampling should yield significantly betters results. Based on the dataset, Can you indicate which variables have a significant effect on the likelihood of heart disease? Do analysis. Lastly deploy the project using gradio.

The dataset contains 18 variables. In machine learning projects, "HeartDisease" can be used as the explonatory variable, but note that the classes are heavily unbalanced.

1. HeartDisease: Respondents that have ever reported having coronary heart disease (CHD) or myocardial infarction (MI)

2. BMI: Body Mass Index (BMI)

3. Smoking: Have you smoked at least 100 cigarettes in your entire

life? [Note: 5 packs = 100 cigarettes]

4. AlcoholDrinking: Heavy drinkers (adult men having more than 14

drinks per week and adult women having more than 7 drinks per we.ek

5. Stroke: (Ever told) (you had) a stroke?

6. PhysicalHealth: Now thinking about your physical health, which includes physical illness and injury, for how many days during the

past 30 days was your physical

7. MentalHealth: Thinking about your mental health, for how many days during the past 30 days was your mental health not good?

8. DiffWalking: Do you have serious difficulty walking or climbing stairs?

9. Sex: Are you male or female?

10. AgeCategory: Fourteen-level age category

11. Race: Imputed race/ethnicity value

12. Diabetic: (Ever told) (you had) diabetes?

13. PhysicalActivity: Adults who reported doing physical activity or exercise during the past 30 days other than their regular job

14. GenHealth: Would you say that in general your health is.

15. SleepTime: On average, how many hours of sleep do you get in a 24-hour period?

16. Asthma: (Ever told) (you had) asthma?

17. Kidney Disease: Not including kidney stones, bladder infection or incontinence, were you ever told you had kidney disease?

18. SkinCancer: (Ever told) (you had) skin cancer?