

1) If you were to only consider the provider's effectiveness with regard to this particular treatment, would you recommend reaching out to any of these doctors in particular? If so, which ones, and why?

Yes. I would recommend patients to reach out doctors: Octopus, Dre, Lower, Faustus, Demento as my top 5 effective doctors.

I assumed missing data in outcome column means the patient is 'cured'. I built a logistic regression using variables: health risk assessment and doctors' name (categorical variable) as my feature.

I check the summary statistics of the logistic regression model. Those doctors with negative coefficients indicate that the failures of their patients are less likely to happen compared with doctor Andrews(my baseline). Those doctors with positive coefficients indicate that the failures of their patients are more likely to happen.

I selected top5 with lowest coefficients as doctors with relatively high effectiveness.

2) What other data would be helpful in understanding which doctors to reach out to? What other data would help you evaluate the overall clinical effectiveness of the doctor? How would you use this data?

It would be helpful if I can have data for other treatments those doctors conducted; Hospitals of doctors(maybe some hospital is better than others?); Patients rating of doctors, and more details about patients that could affect their health conditions(for example, take pills on time, avoid certain food, work out etc.)

I can also use data of other treatments results to evaluate overall clinical effectiveness and use online rating to build different metrics other than patients' re-visit or not as evaluation criteria.