ML Task:

For this ML task we first collected the data from the FHIR server using a python script and exported the data onto a csv file. We then cleaned the data inside the csv file and then imported the data into RStudio. In RStudio we split the data into 70% training, 30% testing split and trained the data using the rpart library. After that we predicted the values into the test data and calculated the accuracy by comparing the actual value and the predicted value using a for loop.

Our dataset consists of 9 identifiers: BMI, Triglycerides, Potassium, Glucose, Hemoglobin A1C, Platelet mean average, Age, Gender, Tobacco Smoker.

	A	B	C	D	E	F	G	H	I	J	K	L
1	Patient No	Total Chol	BMI	Triglycerid	Potassium	Glucose	Hemoglob	Platelet	Age	Gender	ActualValu	Tobacco
2	238	171.28	29.41	114.59	4.31	73.19	6.02	10.378	65	male	IsLow	Never smoker
3	1494	185.67	25.7	119.86	5.13	93.12	5.98	11.692	47	male	IsLow	Former smoker
4	3689	175	27.29	146.27	5.13	94.57	5.99	10.571	97	male	IsLow	Former smoker
5	4714	160.8	30.16	148.72	3.74	88.34	6.18	10.379	66	male	IsLow	Former smoker
6	5309	193.23	28.75	190.03	3.78	67.55		12.115	70	male	IsLow	Former smoker
7	6497	162.06	30.08	131.26	5.19	69.15	6.05	11.146	60	male	IsLow	Never smoker
8	15039	179.53	27.63	146.97	4.03	70.85	6	10.931	106	male	IsLow	Former smoker
9	19583	192.06	30.09	137.2	4.87	65.06	6.33	11.555	72	female	IsLow	Never smoker
10	22830	172.86	27.88	122.38	4.55	65.12	5.98	10.474	77	male	IsLow	Former smoker
11	23152	166.82	18.73	143.72	4.7	65.31	6.39	11.948	85	male	IsLow	Never smoker

Figure 1 Sample Of Dataset

The training revealed that the important identifiers to detect high cholesterol were Triglycerides value and glucose value.

We got an accuracy of 0.96.

We used TotalCholesterol >= 200 as the cutoff for high cholesterol based on online sources (Jenna Fletcher, 2020) and also the data itself. As most of the cholesterol readings were between 170-230.

This is the classification tree that we got from training the dataset.

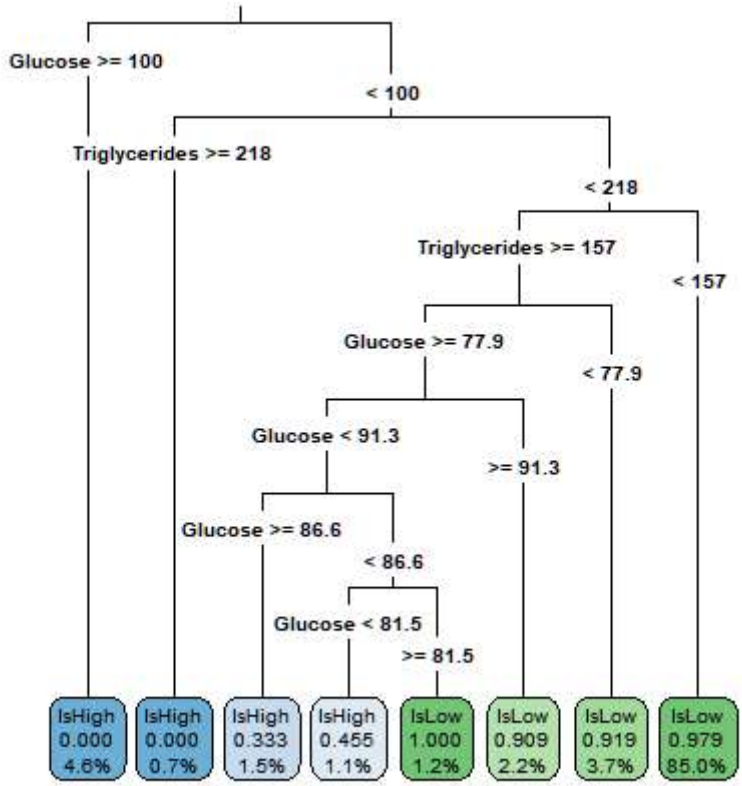

Figure 2 Classification Tree

Resources:

Jenna Fletcher, 2020, https://www.medicalnewstoday.com/articles/315900#tips