

Project Proposal

Group Members: Andres Sanchez Gonzalez, Hiren Jethra, Parthsarthy Nandi

Dataset: <https://archive.ics.uci.edu/ml/datasets/Fertility>

Data Snapshot:

Season	Age	Childish disease	Trauma	Surgical Intervention	High Fevers in Last year	Alcohol Consumption	Smoking Habit	Number of hours spent sitting	Output
-0.33	0.69	0	1	1	0	0.8	0	0.88	N
-0.33	0.94	1	0	1	0	0.8	1	0.31	O
-0.33	0.5	1	0	0	0	1	-1	0.5	N
-0.33	0.75	0	1	1	0	1	-1	0.38	N

Objective:

To use the data to predict seminal quality in males.

Explanatory variables:

1. Season in which the analysis was performed. 1) winter, 2) spring, 3) Summer, 4) fall. (-1, -0.33, 0.33, 1)
2. Age at the time of analysis. 18-36 (0, 1)
3. Childish diseases (ie , chicken pox, measles, mumps, polio) 1) yes, 2) no. (0, 1)
4. Accident or serious trauma 1) yes, 2) no. (0, 1)
5. Surgical intervention 1) yes, 2) no. (0, 1)
6. High fevers in the last year 1) less than three months ago, 2) more than three months ago, 3) no. (-1, 0, 1)
7. Frequency of alcohol consumption 1) several times a day, 2) every day, 3) several times a week, 4) once a week, 5) hardly ever or never (0, 1)
8. Smoking habit 1) never, 2) occasional 3) daily. (-1, 0, 1)
9. Number of hours spent sitting per day ene-16 (0, 1)

Dependent Variable:

10. Output: Diagnosis normal (N), altered (O)