

UNSUPERVISED LEARNING – CASE STUDY: MEDICAL APPOINTMENTS NO-SHOWS

The dataset contains 300k medical appointments described through 14 associated variables. Among these, a feature stating if the patient show-up or no-show to the appointment.

- 'Age': How old is the patient
- 'Gender': Male or Female
- 'AppointmentRegistration': The day someone called or registered the appointment
- 'AppointmentDate': The day of the actual appointment
- 'DayOfTheWeek': The day in the week of the actual appointment
- 'Status': Status of the patient in terms of 'show-up' or 'no-show' (30% are no shows)
- 'Diabetes': True or False
- 'Alcoholism': True or False
- 'HiperTension': True or False
- 'Disability': True or False (original attribute: categorical with 4 values)
- 'Smokes': True or False
- 'Tuberculosis': True or False
- 'Sms_Reminder': 1 or more messages sent to the patient
- 'AwaitingTime ': Number of days between AppointmentRegistration and ApointmentData

Objective: Extracting the most relevant variables affecting no-shows.