**HOSPITAL APPOINTMENT BOOKING**

**PROBLEM STATEMENT**

All over the globe, significant amounts of patients miss their appointments without cancelling in time or even cancelling at all, resulting in billions of dollars wasted yearly due to increased idle time, overtime and waiting time that the other patients and hospitals face.

Hospitals are actively trying to implement methods to try to reduce the idle time caused by patient no-shows by using overbooking and reminder systems.

However, these two methods can be very costly. Overbooking can lead to patient dissatisfaction and constant personalized reminders, such as phone calls, to every patient can be annoying and costly in terms of manpower. This paper focuses on offering a solution which mitigates the global phenomenon of medical no-shows by creating a machine learning model using existing patient datasets to discover patterns and relationships between multiple patient variables and their tendency to miss appointments. Therefore, the likelihood of a patient showing up, given their information, may be predicted.

The machine learning model used to form the solution predictive model is based on the decision tree classification algorithm. Furthermore, a scheduling system was implemented such that the overall model detects whether a patient has a risk of missing an appointment with a 95% accuracy, upon which it automatically enables the risky patient’s schedule slot for overbooking and notifies medical staff or administration to contact them accordingly.

A person makes a doctor appointment, receives all the instructions and no-show. Who to blame?

This dataset collects information from 100k medical appointments and is focused on the question of whether or not patients show up for their appointment.

A number of characteristics about the patient are included in each row.

**DATASET DESCRIPTION**

**PatientId:** Identification of a patient**.**

**AppointmentID:** Identification of each appointment.

**Gender:** Male or Female.

**AppointmentDay**: The day of the actual appointment, when they have to visit the doctor.

**ScheduledDay**: The day someone called or registered the appointment, this is before appointment of course.

**Age:** How old is the patient.

**Neighbourhood**: Where the appointment takes place.

**Scholarship:** True of False, indicates whether or not the patient is enrolled in some welfare program

**Hypertension:** True or False.

**Diabetes:** True or False.

**Alcoholism:** True or False.

**Handicap:** True or False.

**SMS\_received:** 1 or more messages sent to the patient.

No-show

Here, "No Show" means absent. So, "No Show" = "No" means patient attended appointment and "No Show" = "Yes" means patient missed appointment.

EDA Questions

Q1: How often do men go to hospitals compared to women? Which of them is more likely to show up?

Q2: Does receiving an SMS as a reminder affect whether or not a patient may show up? is it correlated with number of days before the appointment?

Q3: Does having a scholarship affects showing up on a hospital appointment? What are the age groups affected by this?

Q4: Does having certain diseases affect whether or not a patient may show up to their appointment? is it affected by gender?