Medical No Show

Kathleen Zhu, Alice Li

September 22, 2017

1 OBJECTIVE

To predict the likelihood a given patient will fail to make a scheduled medical appointment.

2 DATA

We will be using a dataset found on kaggle:
https://www.kaggle.com/joniarroba/noshowappointments/data

This dataset contains 300k historical observations of appointments made in the Brazilian public medical sector. For each observation we have information on 15 features regarding both the appointment (day of week, time, etc.) as well as the patient (age, gender, various health conditions), and whether or not they showed up for their appointment.

3 MEMO

The rising financial burden imposed by health care is a critical social issue that needs to be addressed. These medical costs are caused by a multitude of reasons; however, we do not want to contribute to these costs through trivial administrative wastes accumulated through patients missing appointments. This results in both unnecessary costs to patients as well as time wasted for doctors.

We hope to reduce these inefficiencies by predicting which appointments and patients are likely to be no shows. The information contained in the dataset will help us examine which features are associated with the missed appointments. We can aggregate the data at different levels of granularity looking at features such as the time of the appointment, or the age of the patient to start unpacking relationships between the variables.

Our efforts in creating this predictive model will help optimize the appointment process. We will at once be more time-efficient as well as cost-effective.