Predicting “No Show” for Medical Appointment

Presented by : Meng

Temasek Polytechnic, School of Informatics & IT – Up-Skill in Business Analytics

INTRODUCTION & OBJECTIVES

INTRODUCTION

Governments are spending millions each year to improve healthcare and healthcare spending is increasing year over year, Brazil is no exception.

Being a nation that has huge population, limited finances, limited healthcare capacity and high demand for healthcare services, such wastage should be reduced.

As part of preventive healthcare, patients should show up for their appointments for diagnosing and

Should a patient misses, their appointment may lead to deterioration of their illness and the need of more critical care in future.

OBJECTIVE :

To predict the “No Show” rate of medical appointments and recommend ways to improve show up rate.

DATA EXPLORATION and DATA PREPARATION

DATA EXPLORATON

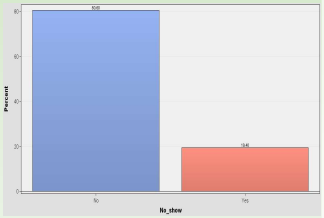
The dataset consists of 110,527 rows and 14 columns.

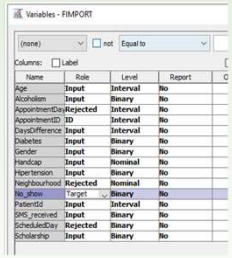
DATA PREPARATION

The columns - PatientID & Neighbourhood are not applicable and will be rejected.

Columns like ScheduledDay(day of scheduling appointment) and AppointmentDay(day of appointment) will not be directly used. Instead, DaysDifference was calculated to using the difference between the 2 dates and injected into the training models.

Preliminary, shows the target variable “No_Show”, No = 80.60% and Yes = 19.40%.





Target variable

“No_show” of was set to “binary”.

The “No” and “Yes” was replaced as “0” and “1” respectively in the replacement node.

“No_Show, NO” = showed up for appointment

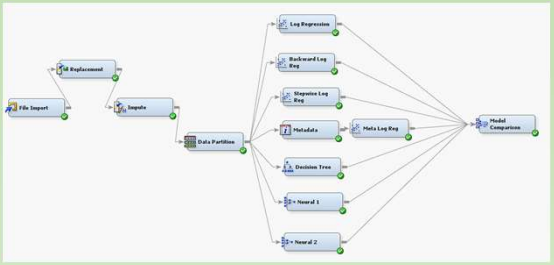
“No_Show, YES” = did NOT show up for appointment

No_show	No	0
No_show	Yes	1

DATA PARTITION

Partition	Allocation	Observed	Function
Training	40%	44,210	Training the model
Validation	30%	33,157	Evaluating the best prediction model
Test	30%	33,160	Testing the implementation model

WORKFLOW



Model comparison was used to determine the best model and comparing the results. With the target variable = REP_No_Show and the Selection Criterion of “Valid Misclassification Rate”

4 models of Logistic Regression was used, from Default Logistic Regression, Stepwise, Backward and a Logistic Regression that was ran through Metadata.

Decision Tree with the following criterion setting for Interval “ProbF”, Nominal – “Entropy”, Oridinal – “Entropy”.

Neural Network, 2 models was imported, Neural 1, with “Model selection Criterion” as “Average Error”, with 3 hidden units in Network and Max of 50 iterations in Optimization

Neural 1, with “Model selection Criterion” set as “Misclassification”, with 24 hidden units in Network and Max of 150 iterations in Optimization

FINDINGS & RECOMMENDATIONS

Data Role=VALIDATE Target Variable=REP_No_show Target Label=Replacement: No-show

Target	Outcome	Target Percentage	Outcome Percentage	Frequency Count	Total Percentage
0	0	79.9739	99.7430	26393	79.6001
1	0	20.0261	98.7007	6609	19.9324
0	1	43.8710	0.2570	68	0.2051
1	1	56.1290	1.2993	87	0.2624

Data Role=VALIDATE Target=REP_No_show Target Label=Replacement: No-show

False Negative	True Negative	False Positive	True Positive
6609	26393	68	87

Actual / Predicted	Predicted - Yes (1)	Predicted - No (0)	Total
Actual Yes (1) No - Show	87	6609	6696
Actual NO (0) No - Show	68	26393	26461
	155	33002	33157

The Validation data as shown :

False Positive (FP) : 0.205%
False Negative (FN) : 19.95%
True Positive (TP): 0.262%
True Negative (TN): 79.600%

Misclassification (Error) Rate : 20.138%
Recall (Sensitivity) : 1.299%
Precision : 56.129%
Accuracy : 79.862%

Since the objective is to predict if the “No Show” rate.

The model, is not that accurate (sensitivity) at 1.299%.

Although the model may have a decent Accuracy rate of 79.862% and fairly decent Misclassification rate.

However, the Sensitivity is very low at 1.299% and is only able to predict correctly at about 56.129% of the time.

As such the data set is not sensitive and further testing of model will not yield improvements in results.

The following are some suggestions to boost show up rate :

1. Impose an upfront fee for medical appointment, which will be offset when showing up for the appointment.
2. Use the patient’s preferred method of reminders to increase show up rate.
3. Educate patient on cost of no show and importance of showing up for appointments.
4. Have a waiting list of appointments, where the waiting list will fill in those who fail to show up.

References :

https://wellapp.com/blog/patient-no-show-rates/
https://www.solutionreach.com/blog/10-truly-awesome-ways-to-reduce-no-shows?hs_amp=true