

# SAS<sup>®</sup> GLOBAL FORUM 2018

---

## USERS PROGRAM

### Medical Appointments: Show/No-Show Prediction using Data Mining

April 8 - 11 | Denver, CO  
**#SASGF**

# Medical Appointments: Show/No-Show Prediction using Data Mining

Shubham Panat

Oklahoma State University

## ABSTRACT

Many times people do not show up for a medical appointment. Previous studies have shown that about 25% of the people did not show up.

No-show is a loss for doctors since they lose their payments. On the other hand, patients who wanted an appointment as soon as possible were unable to get one.

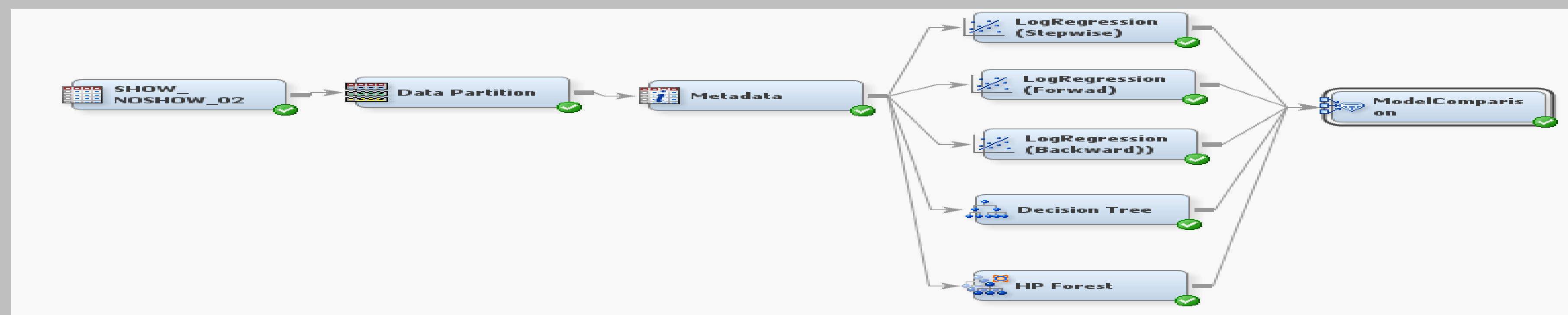
Thus, there are two losses: the financial loss for the doctor and the loss of an appointment for the person in need.

The paper could help clinics and hospitals in understanding what attributes are associated with the individuals who did not show up.

Using SAS® Enterprise Miner™, we try to predict what factors are responsible for a No Show. The paper examines the ways by which we can reduce no-shows. The data is from the Kaggle website and contains 300,000 medical appointments and 15 variables.

## METHODS

- Missing information in the data set was handled.
- Data issues, such as age less than 0, appointment date prior scheduled date e.t.c., were corrected.
- Dummy variables were created (0=Negative, 1=Positive) 'Show\_Status' (1= Show, 0= No Show) is our target variable in the analysis
- After Data Cleaning, Data was partitioned into Training and Validation (70% Train, 30% Validation)
- Model was build on the Training Data Set and Validated on the Validation Data Set.
- Since target is categorical, so Logistic Regression and Decision Tree Classifier were used.
- Model Comparison was performed to choose the best model.



## RESULTS (CLICK TO EDIT)

Name	Use	Report	Role	Level
Age	Default	No	Input	Interval
Alcoholism	Default	No	Input	Nominal
Appointment	No	No	Input	Interval
Diabetes nev	Default	No	Input	Nominal
Gender	No	No	Input	Nominal
Gender cat	Default	No	Input	Nominal
Hypertension	Default	No	Input	Nominal
Neighbourho	Default	No	Input	Nominal
No show	No	No	Input	Nominal
PatientId	No	No	Input	Interval
SMS received	Default	No	Input	Nominal
Scholarship	Default	No	Input	Nominal
Show Status	Yes	No	Target	Nominal
Waiting Days	Default	No	Input	Interval
Week Day	Default	No	Input	Nominal
WeekofDay	No	No	Input	Nominal
handicap	Default	No	Input	Nominal

Target Variable: Show Status

Independent Variable: Age, Alcoholism, Diabetes, Gender\_cat, Hypertension, Neighborhood, SMS\_received, Scholarship, Waiting Days, Week Day, and handicap.

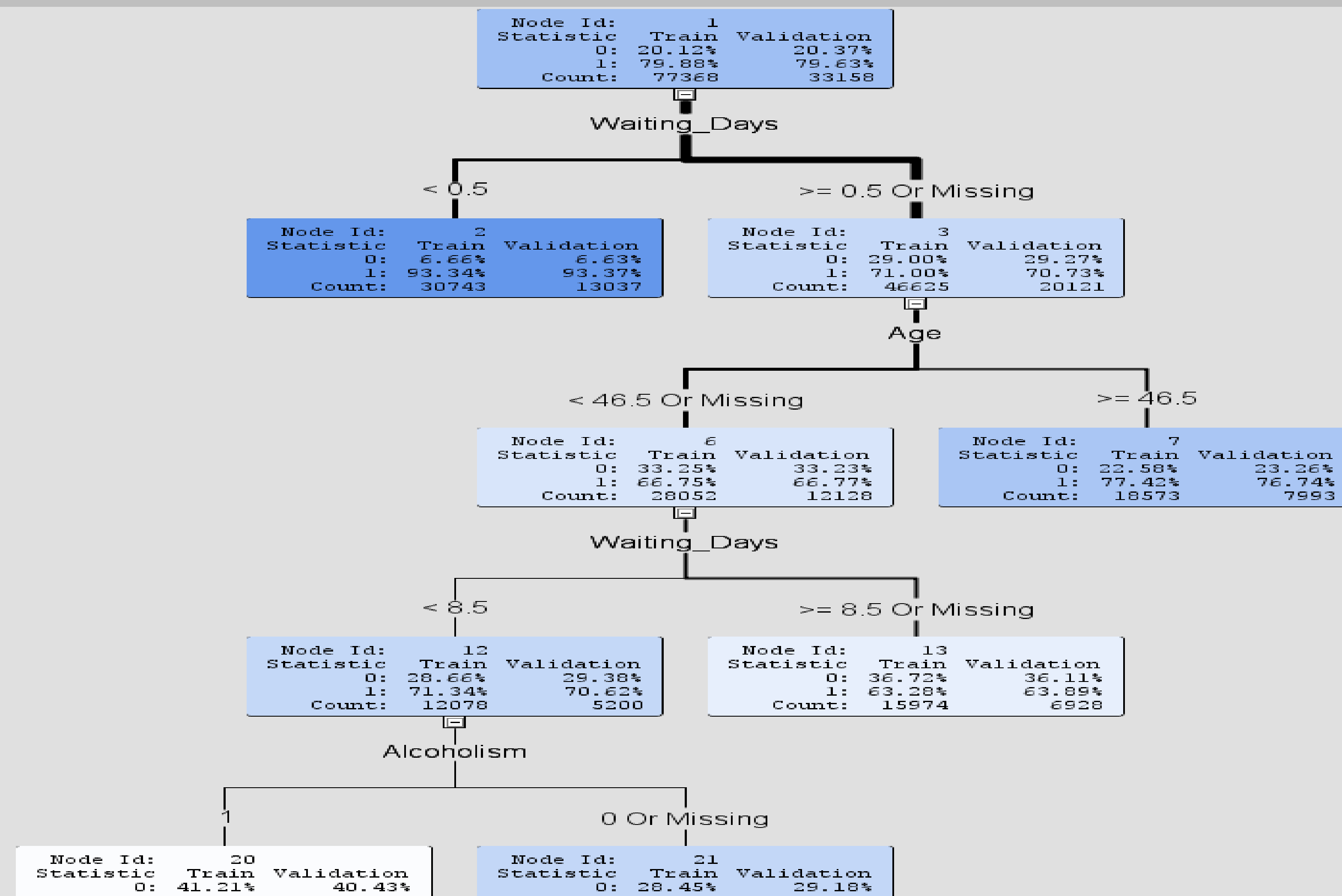
Selected Model	Predecessor Node	Model Node	Model Description	Target Variable	Target Label	Selection Criterion : Valid: Misclassification Rate
Y	Tree	Tree	Decisio...	Show ...		0.20345
	HPDM...	HPDM...	HP For...	Show ...		0.2036...
	Req	Req	LoqRe...	Show ...		0.2068...
	Req2	Req2	LoqRe...	Show ...		0.2068...
	Req3	Req3	LoqRe...	Show ...		0.2068...

# Medical Appointments: Show/No-Show Prediction using Data Mining

Shubham Panat

Oklahoma State University

## RESULTS CONTINUED (CLICK TO EDIT)



## CONCLUSIONS

People who are 46+ generally show up for the appointment.

When Age <46 and Waiting Day is more than 8 days, there are 36% chances of No-Show.

When Age <46, Waiting Day is more than 8 days, and the person is alcoholic, there are 40% chance of No-Show.

When it is the same day appointment, 93% show up.

### Recommendations:

Since waiting days plays a crucial role in the show/no-show, Hospitals should try to minimize this waiting period.

Instead of first come first serve basis, Hospitals may optimize by reserving 60% of appointment on a day to folks who are younger than 45 years and rest 40% for folks above 45 years in age.

It seems that SMS messaging doesn't impact people to show up for their appointment. So Hospitals could simply stop sending SMS, and not incur the cost. The no-show rate is almost the same with or without a reminder.

It would have been great if the data set had the record such as:

Who scheduled the appointment?

Was this the first appointment?

How far was the patient from the facility location?

## REFERENCES

<https://www.kaggle.com/joniarroba/noshowappointments>

[https://www.sas.com/content/dam/SAS/en\\_ca/User%20Group%20Presentations/Winnipeg-User-Group/Rothman-PredictiveModelingSept14.pdf](https://www.sas.com/content/dam/SAS/en_ca/User%20Group%20Presentations/Winnipeg-User-Group/Rothman-PredictiveModelingSept14.pdf)

[http://support.sas.com/documentation/cdl/en/stathpug/68163/HTML/default/viewer.htm#stathpug\\_hpsplit\\_details01.htm](http://support.sas.com/documentation/cdl/en/stathpug/68163/HTML/default/viewer.htm#stathpug_hpsplit_details01.htm)



# SAS<sup>®</sup> GLOBAL FORUM 2018

April 8 - 11 | Denver, CO  
Colorado Convention Center

#SASGF