University of Niš
Faculty of Electronic Engineering

Laboratory for
Medical Informatics

OPTIMIZATION OF HEALTH SERVICE SCHEDULE

- Anđelija Đorđević
- Aleksandar Milenković
- Dragan Janković
- Lazar Stamenković

The content

- Motivation
- Methodology
- Solution
- Further work
- Conclusion

Motivation

- Long waiting time for scheduled medical examinations
- Frequent case of cancelation or absence of patients on scheduled medical care
- The problem of cancellation and absence of scheduled appointments is reduced by predicting whether or not the patient will show up on a scheduled appointment

Year	2012	2013	2014	2015	2016	2017	2018
The proportion of patients who canceled or failed to appear	57%	35%	33%	34%	32%	25%	21%

Methodology

- Supervised learning - logistic regression model with L1 regularization
- The probability of a patient missing the appointment
- The most important attributes for patients' arrival and non-arrival
- Data processing

Input data	Final data
identifier of the insured	*probability of patient cancelation*
start date time	• *day of the month* • *month* • *Serbian feast* • *Christmas and New Year* • *The national holiday* • *day of the week* • *time of the day*
end date time	–
identifier of the doctor	*probability of doctor cancelation*
identifier of the operator	*probability of operator cancelation*
identifier of the operator who canceled the term	–

Input data	Final data
date and time when the appointment is scheduled	• *day of the month* • *month* • *Serbian feast* • *Christmas and New Year* • *The national holiday* • *day of the week* • *time of the day* • *number of waiting days*
age	*age*
gender	*gender*
insurance	*insurance*
status	*label*
specialization	*specialization*
chronic diagnosis	*chronic diagnosis*

Solution

- The created model has an accuracy of 80%
- The predictions must be taken with reserve
- It is very important that the model does not produce an additional problem (arrival of two patients at the same appointment)
- Changing the logistic regression threshold results in different values of model metrics

Threshold	Accuracy (%)	F-measure (%)	precision (%)	recall (%)	sensitivity (%)	specificity (%)	TP	FP	TN	FN
0.1	50.7	57.33	41.26	93.88	93.88	27.18	215626	306954	114568	14055
0.2	67.42	64.16	52.42	82.67	82.67	59.11	189874	172360	249162	39807
0.3	74.82	67.44	62	73.93	73.93	75.31	169814	104094	317428	59867
0.4	78.26	68.61	69.9	67.36	67.36	84.2	154713	66615	354907	74968
0.5	80	68.72	76.69	62.25	62.25	89.69	142968	43448	378074	86713
0.6	80.95	67.89	83.74	57.08	57.08	93.96	131112	25457	396065	98569
0.7	80.76	64.94	90.87	50.52	50.52	97.23	116033	11657	409865	113648
0.8	77.23	53.39	96.71	36.87	36.87	99.32	84692	2877	418645	144989
0.9	67.98	16.98	99.26	9.28	9.28	99.96	21321	158	421364	208360

Which threshold value will be taken depends on whether in a particular situation a larger problem is caused by empty or duplicate appointments

Solution

- The attributes that are most relevant to patients' failure were identified
- Influencing these attributes can lead to a reduction in blank terms
  - National holidays
  - Specializations of doctors: oral surgery, occupational medicine, speech therapist, dental prosthetics, epidemiology and radiology

Further work

- Collecting new data
- Working with other algorithms
- The considered problem of cancellation and non-attendance is not unique and exists in all areas of life where reservations are possible (hotels, restaurants, airlines)
- Lufthansa Airlines uses a similar system

Conclusion

- The implementation would contribute to:
  - improving the schedule
  - reducing the costs of non-attendance of patients
  - reducing patients' waiting times for certain examinations
- The predictions must be taken with reserve
- The beginning of optimization of the health service schedule

Any questions?

Thank you for your attention!