Einar Arvizu Gil (A01350870)
Tecnológico de Monterrey, Campus Querétaro
A01350870@itesm.mx

May 30, 2021

Abstract

Periodontal disease is a chronic infection of the gums characterized by a loss of attachment between the tooth and bone, and by bone loss. In this investigation we are looking to predict how severe periodontitis can be in any given patient by applying a linear regression model. In a multivariate regression controlled by age, diabetes, hypertension, bacterial plaque and bleeding.

Introduction

Periodontitis is a major public health problem due to its high prevalence, as well as because it may lead to tooth loss and disability, negatively affect chewing function and aesthetics, be a source of social inequality, and impair quality of life. Periodontitis accounts for a substantial proportion of edentulism and masticatory dysfunction, results in significant dental care costs and has a plausible negative impact on general health.

Classification of periodontitis based on stages defined by severity (according to the level of interdental clinical attachment loss, radiographic bone loss and tooth loss), complexity and extent and distribution:

Perio Stage	Stage I	Stage II	Stage III	Stage IV
CAL at site	1-2 mm	3-4 mm	>5mm	>5mm
Tooth loss	No tooth loss		Loss <4 teeth	Loss >5 teeth

To find the stage where a patient is, the dentist has to fill and complete a periodontal exam by measuring the interdental CAL in each tooth. And sum up the number of teeth that are in each classification. Finally a rule of three is made to get the percentage of the tooth in each state, the stage with the biggest amount is the one selected for the patient, even if there is some amount in the other stages.

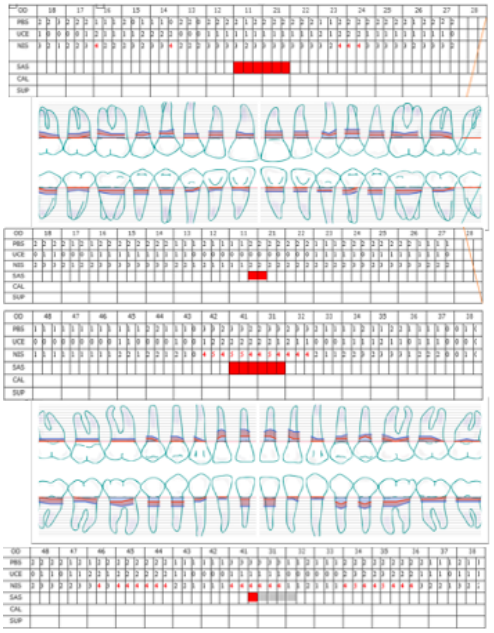


Figura 7. Examen periodontal

Development

A linear regression model can be used to predict the severity of periodontitis in a given patient. We are looking to make this prediction before a dentist has to fill a periodontal exam. This would mean to look for the variables correlated to the disease and that can be easily obtainable.

With multiple results we already can find that there are some diseases that can affect a rapid progression of periodontitis in a patient like diabetes and hypertension.

We are also going to include the age of the patient which has already proved that by becoming older the periodontitis is more common.

And finally we are looking to use the bacterial plaque and bleeding as main factors. In the bacterial plaque we can find microorganisms that favors the formation of caries and periodontal disease, like the loss of tooth. Therefore this exam is vital to identify the level of risk.

With the help of DentoMedica Gil & Brawn we were able to collect the data and evaluate more than 200 people. We had to exclude at least 50 cases where data was incomplete in the percentages of CAL in each stage or lack of information if they had a systemic disease.

After data was collected successfully, we found a way to measure the severity of periodontitis by giving percentages to each stage and transforming the biggest amount of each patient to this new variable. Which will be treated as our independent variable for the linear regression.

Perio Stage	Stage I	Stage II	Stage III	Stage IV
Severity %	16%	32%	32%	20%

To perform the linear regression it was used python and numpy to recreate the linear regression formula, by following these steps:

1. Get the Y predictions for X with the current weight and bias.
2. Calculate the loss from the prediction and targets in Y.
3. Find the gradients with respect to the parameters of weight and bias.
4. Update the values of the weight and bias.

After fitting the model the user can input the variables to predict the percentage of severity. The following table represents how to read the results:

Perio Stage	Stage I	Stage II	Stage III	Stage IV
Severity %	0-16%	17-48 %	49-80 %	81-100 %

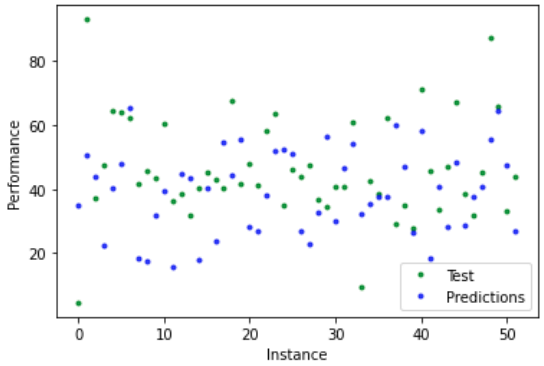
Results

To see how well the variables were related to the severity, a correlation function was executed and it was obtained the following results:

Variable	Correlation
AGE	0.088065
DIABETES	-0.025629
HYPERTENSION	0.094906
STAGE I	-0.511674
STAGE II	-0.317266
STAGE III	0.347761
STAGE IV	0.712325
%Severity	1.000000
%BACTERIAL PLAQUE	0.079297
%BLEEDING	0.210555

At first we were hoping that the severity had a lot more correlation within the plaque, as we can observe the severity has the most correlation with the different stages because it was only formed by them. But each stage was highly related to the plaque and bleeding.

It was also obtained the difference between the test and predictions of the model, we can observe in the graph the variance between the two of them.



This model has a console to make different predictions, and we can see its good functionality by input probable outcomes, like someone with low risk. To actually see how well the model behaves it would be necessary to compare the results to a periodontal exam.

Conclusion

The model was successfully trained, although as we can see there is a mean square error of almost 20. This would mean that on each prediction, the real value can be +15% or - 15%. For this model this error is too high for the severity might change from stage to stage if the error is considered in the prediction.

It is also important to notice that the data used has 170 instances and there are a lot of missed cases having a big impact on the functionality and the precision of the linear regression. As well, there are very few cases for some stages to learn from.

References

Chaple, A. (May, 2019). The O'Leary index, a.k.a. the "Love" index. Cuban Magazine of Stomatology.

Izoura, K et Al. (June, 2016). Dental loss among ambulatory patients with diabetes. Journal of Clinical & Translational Endocrinology

Papanou, P et Al. (March, 2018). Periodontitis: Consensus report of workgroup 2 of the 2017 . Wiley Journal of Clinical Periodontology

Wojtkowska, A. et al. (January, 2021) The inflammation link between periodontal disease and coronary atherosclerosis. BMC Oral Health