

PREDICCIÓN DEL RESULTADO DE OSEOINTEGRACIÓN EN IMPLANTES DENTALES MEDIANTE MÚLTIPLES CLASIFICADORES

Nancy B. Ganz^{(1,2,*), Alicia E. Ares^(1,2) y Horacio D. Kuna⁽²⁾}

⁽¹⁾ Instituto de Materiales de Misiones (IMAM-CONICET). Félix de Azara 1552, N3300LQH, Posadas, Misiones, Argentina.

⁽²⁾ Facultad de Ciencias Exactas, Químicas y Naturales (FCEQyN). Universidad Nacional de Misiones (UNaM). Félix de Azara 1552, N3300LQH, Posadas, Misiones, Argentina.

*E-mail: nancy.bea.ganz@gamil.com

RESUMEN

El campo de la ciencia de datos ha tenido muchos avances respecto a la aplicación y desarrollo de técnicas en el sector de la salud. Estos avances se ven reflejados en la predicción de enfermedades, clasificación de imágenes, identificación y reducción de riesgos, así como muchos otros. Sin embargo, no se han encontrado trabajos que apliquen procesos o técnicas de aprendizaje automático específicamente a la predicción de fracasos en implantes dentales. Este trabajo tiene por objetivo investigar el beneficio de la utilización de múltiples algoritmos de clasificación, para la predicción de fracasos en implantes dentales de la provincia de Misiones, Argentina. La experimentación es realizada con tres conjuntos de datos, un conjunto concerniente a historias clínicas de implantes dentales confeccionado para el estudio de caso, un conjunto generado artificialmente y un conjunto obtenido del repositorio de datos kaggle. En comparación con los clasificadores individuales, el enfoque propuesto obtiene el mayor porcentaje de acierto de los fracasos, logrando un 75% de casos correctamente identificados.

Palabras Clave: combinación de clasificadores, clasificación, aprendizaje automático, implantes dentales, predicción.

CONTEXTO

Esta línea de investigación se lleva a cabo dentro del Programa de Materiales y Fisicoquímica (PROMyF) en el Laboratorio de Ciencia de los Materiales del Instituto de Materiales de Misiones (IMAM), de la Facultad en Ciencias

Exactas, Químicas y Naturales (FCEQyN), de la Universidad Nacional de Misiones (UNaM), en el marco de un plan de tesis doctoral, bajo el nombre de “*Predicción de fracasos en Implantes Dentales mediante la aplicación de Ciencia de Datos*”. Está financiado por el Consejo Nacional de Investigaciones Científicas y Técnicas (CONICET) a través de una “Beca Interna Doctoral” otorgada por Resolución D N° 4869. Además, cuenta con el auspicio del Colegio de Odontólogos de la Provincia de Misiones.

1. INTRODUCCIÓN

Existen diversos antecedentes que combinan o integran clasificadores, para mejorar el porcentaje de acierto o incluso para no sesgar la decisión sobre los resultados de un solo clasificador [1]–[9]. Inspirados en las ideas de estos trabajos, proponemos un procedimiento de aprendizaje automático mediante la utilización de múltiples clasificadores para un dominio de poco conocimiento, como es el caso de los implantes dentales.

El conjunto de datos utilizado se basa en historias clínicas de pacientes que se han sometido a procesos quirúrgicos de colocación de implantes dentales en la Provincia de Misiones, Argentina.

Un paso importante en este trabajo de investigación fue la búsqueda de los mejores clasificadores individuales para el problema. Después de investigar los métodos existentes y teniendo en cuenta los enfoques de combinación, proponemos la utilización de cinco clasificadores: Random Forest (RF) [10], C-Support Vector (SVC) [11], K-Nearest Neighbors (KNN) [12], Multinomial Naive

Bayes (MNB) [13] y Multi-layer Perceptron (MLP) [14]. Estos clasificadores obtuvieron en evaluaciones exploratorias el mejor desempeño en comparación con otras combinaciones exploradas, que han incluido diferentes métodos como: Rpart, Ada, Gradient Boosting Machine (GBM) [15] y distintos clasificadores Naive Bayes [13].

El enfoque de integración propuesto para los cinco clasificadores consistió en combinar los resultados de sus predicciones para determinar el grado de pertenencia a la clase. Así, se pretende lograr una mayor precisión de la que cada uno de los clasificadores logra de manera individual para la etiqueta de clase objetivo (fracaso). La integración de los modelos se realiza con el método votación suave ponderada [16], [17]. Este método implica la definición de pesos asignados a cada clasificador, para ello se realiza una búsqueda en cuadrícula para obtener los valores de los pesos que mejor se ajustan a los clasificadores en función de su desempeño.

Para la experimentación y validación del enfoque propuesto, se emplea un conjunto de datos de casos reales de implantes dentales, un conjunto de datos generado artificialmente con la técnica de sobremuestreo de minorías sintéticas (SMOTE) [18], basado en los casos reales de implantes dentales, y un conjunto de datos obtenido del repositorio kaggle (Heart Disease).

2. LÍNEAS DE INVESTIGACIÓN Y DESARROLLO

Esta línea de investigación propone estudiar, evaluar y aplicar distintos algoritmos de clasificación, para predecir los fracasos de un conjunto de datos de implantes dentales de la Provincia de Misiones. A través de la presente línea de investigación, se logró confeccionar el conjunto de datos necesario para proceder al diseño de un procedimiento, que a través de una metodología híbrida, permita identificar los factores de fracaso. Así mismo, se reconoce como objetivo de interés, el estudio de los tipos de tratamientos de superficie (arenado, ataque ácido, tratamiento térmico, etc.) de los implantes

utilizados, para incorporar como característica a evaluar.

3. RESULTADOS OBTENIDOS

Entre los resultados obtenidos, destacamos la creación de un procedimiento que consistió en:

Paso 1. Leer el conjunto de datos.

Paso 2. Seleccionar la característica objetivo para la predicción.

Paso 3. Obtener el subconjunto de características más relevante, a través del método Chi Squared, el cual es un método muy utilizado para la selección de características categóricas [19]–[21]. Este evalúa el valor de una característica calculando el valor de la estadística de X^2 con respecto a la clase. Está dado por: $X^2 = \sum_{i=1}^r \sum_{j=1}^c \frac{(O_{ij}-E_{ij})^2}{E_{ij}}$.

Paso 4. Dividir el subconjunto resultante del paso 3 de forma aleatoria para preservar la distribución de ambas clases en: 70 % para entrenamiento y 30 % para evaluación [22]–[25] (ver tabla 1). Garantizando que todos los casos se encuentren representados en ambos conjuntos.

Paso 5. Para obtener un modelo robusto y optimizar los resultados de los clasificadores, se realizó una búsqueda en cuadrícula para ajustar los hiper parámetros [22], [25], [26]. Esta búsqueda se efectuó con los datos de entrenamiento de cada uno de los conjuntos de datos. Para este proceso se especificó: un espacio de búsqueda, un algoritmo de optimización o ajuste (GridSearchCV) [27], un método de evaluación (validación cruzada de 10 iteraciones) y una medida de rendimiento (precisión de equilibrio).

Paso 6. Entrenar cada clasificador con los valores óptimos de los hiper parámetros hallados en el paso 5.

Paso 7. Realizar la predicción de cada clasificar con los datos de prueba.

Paso 8. Integrar las predicciones mediante la aplicación del método votación suave ponderada. Este método recoge las probabilidades de clase predichas por cada clasificador, multiplica por un peso asignado al

mismo y los promedia. La etiqueta de clase final se deriva de la etiqueta de clase con la probabilidad promedio más alta. Está dado por: $\hat{y} = \arg \max_i \sum_{j=1}^m w_j p_{ij}$. El peso fue determinado mediante una búsqueda en cuadrícula utilizando un parámetro de prueba w con valores comprendidos entre 0 y 1. Esta búsqueda fue sometida a una validación cruzada de 10 iteraciones, donde se medió la precisión de

cada clasificador para la clase en cuestión, seleccionando el valor de w que logró la mejor precisión [7], [9].

Paso 9. Validar los resultados hallados (ver Tabla 2). Los parámetros utilizados para evaluar el rendimiento de los clasificadores individuales y comparar con el enfoque propuesto, fueron: matriz de confusión, sensibilidad, especificidad y precisión [28], [29].

Tabla 1. Características de los conjuntos de datos utilizados para la evaluación experimental. De izquierda a derecha se presenta: nombres de los conjuntos de datos, número de muestras, número de atributos por tupla, cantidad de características seleccionadas por el método Chi-Square (X^2) y tamaño de los conjuntos de entrenamiento y de prueba.

Conjuntos de Datos	Muestra	Características	X^2	Entrenamiento	Prueba
<i>Implantes Dentales</i>	1165	33	17	815	350
<i>Artificial</i>	1748	33	21	1223	525
<i>Problemas Cardíacos</i>	303	13	10	212	91

Tabla 2. Eficiencia en el acierto de los clasificadores RF, SVC, KNN, MNB, MLP y el enfoque propuesto (Integrado) sobre los conjuntos de datos *Implantes Dentales*, *Artificial* y *Problemas Cardíacos*.

Conjuntos de Datos	Clasificadores	Clase objetivo	Clase no-objetivo
		(Fracaso)	(Éxito)
		Sensibilidad	Especificidad
<i>Implantes Dentales</i>	RF	59 %	98 %
	SVC	64 %	99 %
	KNN	64 %	99 %
	MNB	72 %	79 %
	MLP	66 %	97 %
	Integrado	75 %	96 %
<i>Artificial</i>	RF	81 %	97 %
	SVC	81 %	99 %
	KNN	81 %	99 %
	MNB	60 %	81 %
	MLP	82 %	97 %
	Integrado	89 %	97 %
<i>Problemas Cardíacos</i>	RF	81 %	71 %
	SVC	70 %	79 %
	KNN	70 %	76 %
	MNB	77 %	74 %
	MLP	72 %	68 %
	Integrado	94 %	58 %

En la tabla 1, se observa que los clasificadores SVC y KNN lograron el mejor rendimiento sobre la clase no objetivo para todos los conjuntos de datos en comparación con los demás clasificadores, incluso superan al enfoque propuesto en todos los casos. Para la clase objetivo, se aprecia que la integración de las predicciones de los cinco clasificadores permitió alcanzar el mayor porcentaje de acierto. Para esta clase, además se observa que el rendimiento de los clasificadores individuales fue variado. Si bien, el rendimiento de la integración de las predicciones no fue la mejor opción para la clase no objetivo, no quiere decir que haya sido la peor en comparación con las predicciones individuales. Mientras que la integración de las probabilidades para la clase objetivo fue la mejor opción, ya que permitió obtener el mayor porcentaje de acierto.

4. CONCLUSIONES

Se demostró que en este dominio es mejor integrar las predicciones de los clasificadores, para no sesgar la decisión sobre un solo resultado. Asimismo, utilizar predicciones integradas permite conocer diversos puntos de vistas o resultados para un mismo caso, ya que se utiliza más de un clasificador, permitiendo asegurar una asignación de etiqueta o clasificación más precisa.

Como resultado, cada clasificador logró hasta un 72% de acierto de la clase objetivo para el conjunto de implantes dentales, mientras que el enfoque propuesto permitió alcanzar el 75% de casos correctamente identificados como fracasos.

Finalmente, se plantea como trabajo futuro validar el enfoque propuesto con otros conjuntos de datos del área de la salud o la medicina. Además, se propone la inclusión o ampliación de los clasificadores utilizados, para evaluar la posibilidad de ajustar el porcentaje de acierto de ambas clases. Así como, realizar una validación con expertos humanos.

5. FORMACIÓN DE RECURSOS HUMANOS

Este proyecto es parte de las líneas de investigación del “Programa de Materiales y Físicoquímica” de la FCEQyN – UNaM, con cuatro integrantes: un investigador independiente del CONICET, un docente categoría I perteneciente al Depto. de Informática de la FCEQyN – UNaM, un doctorando y un tesista de grado de la carrera de Licenciatura en Sistemas de Información.

6. BIBLIOGRAFÍA

- [1] J. Fierrez, A. Morales, R. Vera-Rodriguez, and D. Camacho, “Multiple classifiers in biometrics. part 1: Fundamentals and review,” *Inf. Fusion*, vol. 44, no. December 2017, pp. 57–64, 2018.
- [2] Y. Miao, H. Jiang, H. Liu, and Y. dong Yao, “An Alzheimers disease related genes identification method based on multiple classifier integration,” *Comput. Methods Programs Biomed.*, vol. 150, pp. 107–115, 2017.
- [3] G. Chandrashekar and F. Sahin, “A survey on feature selection methods,” *Comput. Electr. Eng.*, vol. 40, no. 1, pp. 16–28, 2014.
- [4] C. Catal and M. Nangir, “A sentiment classification model based on multiple classifiers,” *Appl. Soft Comput. J.*, vol. 50, pp. 135–141, 2017.
- [5] M. Pandey and S. Taruna, “Towards the integration of multiple classifier pertaining to the Student’s performance prediction,” *Perspect. Sci.*, vol. 8, pp. 364–366, 2016.
- [6] J. Yan, D. B. Bracewell, F. Ren, and S. Kuroiwa, “Integration of Multiple Classifiers for Chinese Semantic Dependency Analysis,” *Electron. Notes Theor. Comput. Sci.*, vol. 225, no. C, pp. 457–468, 2009.
- [7] D. Ruano-Ordás, I. Yevseyeva, V. B. Fernandes, J. R. Méndez, and M. T. M. Emmerich, “Improving the drug discovery process by using multiple classifier systems,” *Expert Syst. Appl.*, vol. 121, pp. 292–303, 2019.
- [8] L. Oliveira, U. Nunes, and P. Peixoto, “On Exploration of Classifier Ensemble Synergism in Pedestrian Detection,” *IEEE Trans. Intell. Transp. Syst.*, vol. 11, no. 1, pp.

- 16–27, 2010.
- [9] H. F. Nweke, Y. W. Teh, G. Mujtaba, and M. A. Al-garadi, “Data fusion and multiple classifier systems for human activity detection and health monitoring: Review and open research directions,” *Inf. Fusion*, vol. 46, no. June 2018, pp. 147–170, 2019.
- [10] L. Breiman, “Random Forest,” *Mach. Learn.*, vol. 45, no. 1, pp. 5–32, 2001.
- [11] C. Chang and C. Lin, “LIBSVM: A Library for Support Vector Machines,” *ACM Trans. Intell. Syst. Technol.*, vol. 2, no. 3, pp. 1–39, 2011.
- [12] N. S. Altman, “An Introduction to Kernel and Nearest-Neighbor Nonparametric Regression,” *Am. Stat.*, vol. 46, no. 3, pp. 175–185, 1992.
- [13] C. D. Manning, P. Raghavan, and H. Schütze, “Text classification and Naive Bayes,” in *Introduction to Information Retrieval*, Cambridge University Press, 2009, pp. 253–287.
- [14] B. Irie and Sei Miyake, “Capabilities of Three-layered Perceptrons,” *IEEE International Conf. Neural Networks*, pp. 641–648, 1988.
- [15] M. Mohandes, M. Deriche, and S. O. Aliyu, “Classifiers Combination Techniques: A Comprehensive Review,” *IEEE Access*, vol. 6, pp. 19626–19639, 2018.
- [16] L. N. Eeti and K. M. Buddhiraju, “A modified class-specific weighted soft voting for bagging ensemble,” *Int. Geosci. Remote Sens. Symp.*, vol. November, pp. 2622–2625, 2016.
- [17] X. Fan and H. Shin, “Road vanishing point detection using weber adaptive local filter and salient-block-wise weighted soft voting,” *IET Comput. Vis.*, vol. 10, no. 6, pp. 503–512, 2016.
- [18] N. V. Chawla, K. W. Bowyer, L. O. Hall, and W. P. Kegelmeyer, “SMOTE: Synthetic Minority Over-sampling Technique,” *J. Artif. Intell. Res.*, vol. 16, pp. 321–357, 2002.
- [19] X. Jin, A. Xu, R. Bie, and P. Guo, “Machine learning techniques and chi-square feature selection for cancer classification using SAGE gene expression profiles,” *Data Min. Biomed. Appl.*, vol. 3916, pp. 106–115, 2006.
- [20] M. Moran and G. Gordon, “Curious Feature Selection,” *Inf. Sci. (Ny)*, vol. 485, pp. 42–54, 2019.
- [21] I. Sumaiya Thaseen and C. Aswani Kumar, “Intrusion detection model using fusion of chi-square feature selection and multi class SVM,” *J. King Saud Univ. - Comput. Inf. Sci.*, vol. 29, no. 4, pp. 462–472, 2017.
- [22] B. T. Pham, M. D. Nguyen, K. T. T. Bui, I. Prakash, K. Chapi, and D. T. Bui, “A novel artificial intelligence approach based on Multi-layer Perceptron Neural Network and Biogeography-based Optimization for predicting coefficient of consolidation of soil,” *Catena*, vol. 173, no. September 2018, pp. 302–311, 2019.
- [23] G. Isabelle, W. Maharani, and I. Asror, “Analysis on Opinion Mining Using Combining Lexicon-Based Method and Multinomial Naïve Bayes,” *2018 Int. Conf. Ind. Enterp. Syst. Eng. (ICoIESE 2018)*, vol. 2, no. IcoIESE 2018, pp. 214–219, 2019.
- [24] K. Bhattacharjee and M. Pant, “Hybrid Particle Swarm Optimization-Genetic Algorithm trained Multi-Layer Perceptron for Classification of Human Glioma from Molecular Brain Neoplasia Data,” *Cogn. Syst. Res.*, vol. 58, pp. 173–194, 2019.
- [25] D. Chong, N. Zhu, W. Luo, and X. Pan, “Human thermal risk prediction in indoor hyperthermal environments based on random forest,” *Sustain. Cities Soc.*, vol. 49, no. April, p. 101595, 2019.
- [26] S. Xu, “Bayesian Naïve Bayes classifiers to text classification,” *J. Inf. Sci.*, vol. 44, no. 1, pp. 48–59, 2018.
- [27] scikit-learn, “Tuning the hyper-parameters of an estimator,” 2019. [Online]. Available: https://scikit-learn.org/stable/modules/grid_search.html#grid-search. [Accessed: 04-Jul-2019].
- [28] H. He and E. A. Garcia, “Learning from imbalanced data,” *IEEE Trans. Knowl. Data Eng.*, vol. 21, no. 9, pp. 1263–1284, 2009.
- [29] M. Sokolova and G. Lapalme, “A systematic analysis of performance measures for classification tasks,” *Inf. Process. Manag.*, vol. 45, no. 4, pp. 427–437, 2009.