

Clasificadores basados en reglas y selección de atributos para el diagnóstico clínico de subtipos del Síndrome de Guillain-Barré

Daniel Alarcón-Narváez¹, Jerusa Marchi², José Hernández-Torruco¹,
Oscar Chávez-Bosquez¹, Betania Hernández-Ocaña¹

¹ Universidad Juárez Autónoma de Tabasco,
División Académica de Ciencias y Tecnologías de la Información,
México

² Universidade Federal de Santa Catarina,
Departamento de Informática e Estatística,
Brasil

{jose.hernandezt,oscar.chavez,betania.hernandez}@ujat.mx,
182H13002@alumno.ujat.mx, jerusa.marchi@ufsc.br

Resumen. Los algoritmos de aprendizaje automático se han utilizado para identificar diferentes tipos de enfermedades, a través de modelos diagnósticos o predictivos. El síndrome de Guillain-Barré (SGB) es un trastorno neurológico que generalmente afecta a personas de cualquier edad y sexo. En este trabajo, vamos a investigar un modelo de diagnóstico para los subtipos del SGB utilizando exclusivamente variables clínicas. Para ello aplicamos los clasificadores: (1) JRip, (2) OneR y (3) PART, comparándolos entre sí para encontrar el mejor de ellos. También se realizaron experimentos con los filtros: (1) CFS, (2) Chi.Squared, (3) Consistency y (4) RandomForest. Estos experimentos utilizaron un conjunto de datos reales que incluye 129 registros de pacientes mexicanos y 26 variables clínicas representativas. Se registraron las medidas de rendimiento estándar y se utilizó el *accuracy* balanceado para medir el rendimiento de los modelos y obtener el mejor clasificador en cada experimento. Encontramos con estos experimentos que OneR con filtro Chi.squared, así como OneR con filtro RandomForest, obtienen el mejor *accuracy* balanceado con un valor de 0.6564.

Palabras clave: SGB, métodos de selección de atributos, clasificación multiclase, clasificadores individuales, medidas de desempeño.

Rule Based Classifiers and Feature Selection for Clinical Diagnosis of Guillain-Barré Syndrome Subtypes

Abstract. Machine learning algorithms have been used to identify different types of diseases, through diagnostic or predictive models. The

Guillain-Barré syndrome (GBS) is a neurological disorder that generally affects people of any age and sex. In this work, we aim to investigate if it is possible to create a diagnostic model for GBS subtypes using exclusively clinical variables, since according to the reviewed literature, there is no model with these characteristics. For this, we applied the classifiers: (1) JRip, (2) OneR and (3) PART, comparing them to find the best one. We also made experiments with filters: (1) CFS, (2) Chi.Squared, (3) Consistency and (4) RandomForest. These experiments used a real data set that includes 129 mexican patient records and 26 representative clinical variables. Standard performance measures were recorded and the balanced accuracy was used to measure the performance of the models and obtain the best classifier in each experiment. We found with these experiments than OneR with Chi.squared filter, as well as OneR with RandomForest filter, they get the best balanced *accuracy* with a value of 0.6564.

Keywords: SGB, feature selection methods, multiclass classification, single classifiers, performance measures.

1. Introducción

El Síndrome de Guillain-Barré (SGB) es una enfermedad en la que el sistema inmunitario ataca al propio cuerpo por error, confundiendo la mielina con una célula extraña. Causa parálisis flácida, debilidad de músculos y extremidades [17]. Eso se debe comúnmente a la inflamación multifocal de las raíces espinales y nervios periféricos. Este síndrome difiere en términos de su apariencia, duración, simetría de manifestaciones clínicas y si dañan o no principalmente la mielina, axón o fibras nerviosas periféricas. En casos graves, las prolongaciones de las neuronas que son las responsables de enviar el impulso nervioso también se dañan [13]. La incidencia anual estimada del SGB es 0.61-2 casos por 100,000 personas y aproximadamente el 25 % de los pacientes con SGB requieren cuidados intensivos. A pesar de un adecuado tratamiento, el 3.5 % muere debido a complicaciones relacionadas con la parálisis respiratoria, ataque cardíaco o trombosis [28]. Los subtipos más comunes del SGB se dividen en cuatro: polineuropatía desmielinizante inflamatoria aguda (AIDP), neuropatía axonal motora aguda (AMAN), neuropatía axonal sensorial motora aguda (AMSAN) y síndrome de Miller-Fisher (MF) [23]. Debido a la variación en gravedad y tratamiento entre subtipos, la diferenciación entre ellos es crítica.

Existen investigaciones en la literatura, donde se realizaron experimentos para la creación de modelos para el diagnóstico del SGB. Por mencionar algunos, tenemos el trabajo de Hernández-Torruco et al. [9]. En ese proyecto se utilizaron un conjunto de datos que contenía pruebas clínicas, serológicas y de conducción nerviosa y aplicaron el árbol de decisión C4.5, SVM (*Support Vector Machines*) usando un núcleo gaussiano y kNN (*k Nearest Neighbour*); alcanzando de *accuracy* con C4.5 0.9211, con kNN 0.9179 y SVM 0.9154. Canul-Reich et al. [1] crearon diferentes modelos predictivos basados en clasificadores individuales,

utilizando un conjunto de datos con 4 variables clínicas y 12 obtenidas de una prueba de conducción nerviosa. Aplicaron clasificadores con diferentes enfoques como árboles de decisión (C4.5), basados en instancias (kNN), basados en kernel (SVM), redes neuronales como SLP (*Single-Layer Perceptron*, por sus siglas en inglés), MLP (*Multilayer feed-forward*, por sus siglas en inglés) y RBF-DDA (*Radial Basis Function-Dynamic Decay Adjustment*, por sus siglas en inglés), clasificadores basados en reglas (OneR, JRip), Regresión logística y método OvA (One vs All). El mejor *accuracy* alcanzado fue del 0.9235.

Torres-Vásquez et al. [25] utilizaron la técnica SMOTE (*Synthetic Minority Over-sampling Technique*, por sus siglas en inglés) para balancear las instancias de cada subtipo del síndrome en el dataset, con el objetivo de investigar si era posible mejorar los modelos predictivos encontrados anteriormente. Aplicaron los clasificadores C4.5, SVM, JRip y la técnica OvO (*One vs One*) para crear subproblemas binarios de los subtipos del síndrome. Estos resultados se evaluaron usando la métrica AUC (*Area Under the ROC, Receiver Operating Characteristic, Curve*, por sus siglas en inglés). El mejor resultado lo alcanzaron con JRip para los subtipos AMSAN vs MF con un valor de 0.8720 de AUC. En ese trabajo pudieron demostrar que algunos modelos con datos balanceados, superan a los modelos con datos desbalanceados.

También Torres-Vásquez et al. [24] usaron 4 métodos de submuestreo para balancear las instancias, con diferentes enfoques como: *Random Undersampling* (RUS), *Tomek Link* (TML), *One Side Selection* (OSS) y *Neighborhood Cleaning Rule* (NCR). Un método de sobremuestreo fue probado para balancear las instancias de cada subtipo del síndrome. Aplicaron los clasificadores C4.5, SVM, JRip y las técnicas de binarización OvA y OvO. De los experimentos de muestreo que realizaron, se obtuvieron los mejores resultados combinando SMOTE con OvA. Con respecto a los clasificadores, obtuvieron con JRip el mejor rendimiento con un valor mayor a 0.90 de AUC, ya que encontró más casos con diferencias estadísticamente significativas para todos los experimentos. Los resultados que obtuvieron en esa investigación mostraron que balancear el conjunto de datos original mejora los modelos predictivos anteriores.

Todos estos trabajos mencionados anteriormente tienen en común el dataset, donde utilizan solo 4 variables clínicas y 12 variables de pruebas de conducción nerviosa. Estas 16 variables fueron las relevantes encontradas a partir de un dataset de 156 variables, usando QSA-PAM (*Quenching Simulated Annealing-Partitions Around Medoids*, por sus siglas en inglés) [2]. Por lo que el 75% de estas variables requieren de la visita a un laboratorio para realizarte las pruebas, lo que implica pérdida de tiempo. Tomando en cuenta que las primeras semanas de este síndrome son cruciales para el bienestar del paciente en este trabajo decidimos utilizar solo variables clínicas. Es decir variables que se basan en la experiencia del especialista y lo que siente el paciente al momento de ir a consulta. Reduciendo así los tiempos para la detección de alguno de los subtipos del síndrome y poder dar el tratamiento adecuado lo antes posible.

En este estudio utilizamos un conjunto reducido de 25 atributos clínicos, seleccionados de un conjunto de datos original de 365 atributos.

Estos atributos seleccionados corresponden a variables clínicas y fueron tomadas de 129 pacientes mexicanos que contienen los cuatro subtipos de SGB mencionados anteriormente. Para los experimentos, aplicamos clasificadores individuales a través de los algoritmos JRip, OneR y PART. Estos tres clasificadores son basados en reglas [12].

Un clasificador basado en reglas utiliza un conjunto de reglas *IF-THEN* para la clasificación. Una regla *IF-THEN* es una expresión de la forma **IF condition THEN conclusion**. La parte *IF* de una regla se conoce como el antecedente o precondición de la regla. La parte *THEN* es la regla consecuente [8]. El clasificador basado en reglas tiene varias ventajas, es altamente expresivo, fácil de interpretar, fácil de generar, puede clasificar nuevas instancias rápidamente, y el rendimiento es comparable con árboles de decisiones [5]. Los clasificadores basados en reglas tienen dos criterios importantes: el *accuracy* de la clasificación en un conjunto de prueba independiente y la diversidad de reglas en los mismos, lo que los hace interesantes en el campo de aprendizaje automático [11]. Existen algunos indicios en la literatura de que reglas muy simples pueden lograr una precisión sorprendentemente alta en muchos conjuntos de datos [18].

En conclusión, escogimos estos clasificadores basados en reglas porque son fáciles de implementar de manera efectiva. Por el hecho también que son fáciles de interpretar para el ser humano, debido a las reglas que generan [3].

También aplicamos filtros, con el objetivo de seleccionar las variables más importantes para la clasificación. Los filtros evalúan las características intrínsecas de los atributos y en base a ello identifican las variables relevantes para un proceso de clasificación. Además tienen la ventaja de tener un bajo costo computacional. En este estudio seleccionamos cuatro filtros: Chi.squared, CFS (Selección de atributos basada en correlación), Consistency y RandomForest. En cada modelo elaborado, se realizaron 30 ejecuciones con una semilla diferente. Lo cual proporciona variabilidad en las instancias usadas en cada ejecución. Los experimentos se realizaron utilizando el lenguaje R bajo el entorno RStudio 1.1.463. Utilizamos los paquetes *psych* [19], *Fselector* [21], *caret* [16], *pROC* [20], *rJava* [26] y *Rweka* [12].

Este documento se organizó de la siguiente manera: en la sección 2 presentamos una descripción del conjunto de datos, los algoritmos de aprendizaje automático y las medidas de rendimiento utilizadas en el estudio. La sección 3 describe el procedimiento experimental. En la sección 4 presentamos y discutimos los resultados experimentales. Finalmente, en la sección 5 resumimos los resultados y damos conclusiones del estudio.

2. Materiales y métodos

2.1. Conjunto de datos

Los datos utilizados en este trabajo se recopilaron en el Instituto Nacional de Neurología y Neurocirugía desde 1993 hasta 2002. Hay 129 registros de pacientes, cada uno clasificado con un tipo de los subtipos de SGB: 20 pacientes con AIDP, 37 con AMAN, 59 con AMSAN y 13 con Miller-Fisher.

Tabla 1. Conjunto de atributos usados en este trabajo.

Variable	Nombre de la característica
v2	Criterios que cumple para el Diagnóstico
v4	Edad
v7	Días de inicio de disminución de fuerza muscular o compromiso de nervios craneales al evento previo
v9	Días antes de acudir al médico del inicio de los síntomas
v10	Días de inicio de la dificultad respiratoria
v36	Cuantos días requirió de la máquina para respirar
v5	Sexo
v21	Debilidad
v23	Parestesias (cosquilleos, ardor, molestias en piel)
v34	Disnea (dificultad respiratoria)
v35	Requirió VMA (Ventilación Mecánica Asistida)
v6	Patología evento previo
v22	Simetría
v24	Miembros superiores-fuerza muscular (Fuerza en brazos)
v25	Miembros inferiores - fuerza muscular (Fuerza en piernas)
v26	Lugar donde iniciaron los síntomas
v27	Reflejos
v29	Afecta músculos extraoculares
v30	Ptosis (caída de parpado, III nervio craneal)
v31	Afecta cerebelo
v32	Marcha atáxica (falta de control muscular o de coordinación voluntaria "clínico")
v33	Nervios Craneales involucrados
v37	Complicaciones
v38	Involucra esfínter urinario y rectal
vS	Estación (primavera, verano, otoño, invierno)

El conjunto de datos original tiene 365 atributos, de los cuales los primeros 38 se consideran clínicos, los otros 327 atributos corresponden a pruebas de laboratorio, tratamiento y seguimiento del paciente. De los 38 atributos, 13 fueron descartados debido a que pertenecían a datos como el número de caso, el número de archivo, la admisión hospitalaria y la fecha de alta, la duración en días en diversas situaciones, dejando 25 atributos clínicos relevantes que se muestran en la Tabla 1.

2.2. Algoritmos de aprendizaje automático

JRip: Aprendizaje basado en reglas que implementa el algoritmo de poda incremental repetida para la reducción de errores (RIPPER, por las siglas en Inglés de *Repeated Incremental Pruning to Produce Error Reduction*). JRip identifica las clases creando un conjunto de reglas [10]. La regla tiene la forma siguiente:

```
si (atributo1 <operador relacional> valor1 <operador logico>
    atributo2 <operador relacional> valor2 <...>) entonces
    valor-decisión
```

OneR: Es la abreviatura de Una Regla, es un algoritmo individual de clasificación de aprendizaje automático que genera un árbol de decisión de un nivel. OneR puede inferir reglas de clasificación típicamente simples a partir de un conjunto de instancias que son sencillas para interpretar. Este algoritmo crea una regla para cada atributo en los datos de entrenamiento y luego selecciona la mejor regla como su única regla [4].

PART: Usa el enfoque de Frank y Witten [6], la idea clave en este clasificador es construir un árbol de decisión parcial en lugar de explorar uno completo. Una de las mayores ventajas de PART sobre otros métodos es la simplicidad, combinando dos paradigmas de aprendizaje de reglas produciendo un buen conjunto de reglas sin la necesidad de optimización global.

2.3. Selección de atributos

Los filtros son técnicas de selección de atributos que reducen o establecen atributos por nivel de prioridad del conjunto de datos para tratar de aumentar el rendimiento [15]. Se tomaron del paquete *Fselector* cuatro filtros: CFS, Consistency (usando medida de consistencia), Chi.squared (basado en una prueba de chi-cuadrado) y RandomForest (usando el algoritmo RandomForest). Los dos primeros son algoritmos que encuentran un subconjunto de atributos para datos continuos y discretos; los dos restantes encuentran pesos de atributos discretos [21].

2.4. Medidas de rendimiento

Las medidas de rendimiento son técnicas estadísticas utilizadas para evaluar la calidad del modelo creado.

Se obtuvieron y registraron las medidas de rendimiento clásicas de la literatura [14]. Nuestra medida de rendimiento principal de los modelos fue el *accuracy* balanceado, pues como se indica debajo, se utiliza para cuando los datos no están balanceados. Tal es el caso de nuestro dataset.

Accuracy: Mide el porcentaje de respuestas correctas del modelo [22]:

$$Accuracy = \frac{TP + TN}{TP + TN + FP + FN}, \quad (1)$$

donde: TP = Verdadero positivo, TN = Verdadero negativo, FP = Falso positivo, FN = Falso negativo.

Accuracy Balanceado: Es la media aritmética de sensibilidad y especificidad, o la exactitud promedio obtenida en cualquiera de los subtipos del SGB cuando los datos no están balanceados [22,27]:

$$AccuracyBalanceado = \frac{\left(\frac{TP}{(TP+FN)}\right) + \left(\frac{TN}{(FP+TN)}\right)}{2}. \quad (2)$$

Sensibilidad: Se puede definir como la capacidad de la prueba para clasificar correctamente los subtipos del SGB [22]:

$$Sensibilidad = \frac{TP}{TP + FN}. \quad (3)$$

Especificidad: Es el complemento de la especificidad, se conoce como falsos positivos [22]:

$$Especificidad = \frac{TN}{TN + FP}. \quad (4)$$

Área bajo la curva ROC: La métrica AUC es un gráfico bidimensional que proporciona un resumen del rendimiento de un modelo de clasificación en presencia de conjuntos de datos desequilibrados con error desigual. Se utiliza una medida llamada AUC, que varía de 0 a 1 (que representa el verdadero positivo y el falso positivo) asociada con la curva ROC. [7].

2.5. Validación

Utilizamos la evaluación *train-test* para los experimentos, con dos terceras partes de los datos para el entrenamiento y una tercera parte para pruebas.

3. Diseño experimental

En la sección 2.1 se describe el conjunto de datos utilizado para el diseño de los modelos. Primero se crearon tres modelos usando todas las variables del conjunto de datos con los clasificadores JRip, OneR y PART. Realizamos 30 ejecuciones independientes, en cada ejecución calculamos el *Accuracy* balanceado.

Al final calculamos el promedio de las 30 ejecuciones del *Accuracy* balanceado, la desviación estándar e identificamos el mejor y el peor valor de las ejecuciones efectuadas.

Siguiente a la creación de los modelos anteriores, realizamos la selección de los atributos mediante los filtros seleccionados, obteniendo un total de 12 modelos diferentes, un modelo por filtro aplicado a cada clasificador. El *accuracy* balanceado es la métrica principal en cada modelo.

Los filtros CFS y Consistency proporcionan una lista de los atributos más relevantes. Con ellos creamos un subconjunto a partir del conjunto de datos para

realizar los experimentos con los clasificadores, tal como se describe en el primer párrafo de esta sección.

Los filtros Chi.squared y RandomForest asignan una puntuación a cada atributo, que significa su relevancia. De cada filtro, tomamos los dos mejores atributos, los mejores tres atributos y así sucesivamente hasta tomarlos todos. Con cada uno de estos atributos, creamos conjuntos de datos que contienen las instancias del SGB junto con el atributo clase. La idea fue evaluar los modelos de diagnóstico que contenían los atributos más relevantes, de acuerdo a cada uno de estos filtros.

4. Resultados y discusión

La Tabla 2 muestra los subconjuntos seleccionados por los filtros CFS y Consistency, también muestra los atributos con los que se alcanzaron los mejores resultados: para Chi.squared el mejor resultado se logró con 5 atributos y para RandomForest con 4 atributos. Se puede observar que la variable v25 (que hace referencia a la falta de fuerza en las extremidades inferiores) y la variable v29 (referente a la afectación de los musculos que mueven el ojo), fueron seleccionadas por los cuatro filtros, por lo que se infiere la relevancia de las variables para los modelos.

Tabla 2. Atributos seleccionados por cada filtro.

CFS	Consistency	Chi-squared	RandomForest
v2	v6	v21	v25
v21	v25	v2	v30
v22	v26	v30	v29
v24	v27	v25	v2
v25	v29	v29	-
v29	v35	-	-
v30	v8	-	-
v31	-	-	-
v33	-	-	-

La Tabla 3 muestra los resultados de los experimentos con todas las variables del conjunto de datos, ordenados de la forma en como se explica en la sección 3. En negritas se muestra el resultado más alto logrado por uno de los tres clasificadores.

La Tabla 4 muestra la comparación de los mejores resultados obtenidos por los clasificadores JRip, OneR y PART con cada filtro correspondiente, así como el promedio de las 30 ejecuciones, la desviación estándar, el mejor y el peor resultado. Se puede observar por medio de las tablas que de manera general los filtros mejoran el rendimiento de los clasificadores. El resultado promedio más alto con todas las variables del conjunto de datos fue de 0.6265 y los resultados

Tabla 3. Rendimiento de los clasificadores basados en reglas, usando todas las variables en el dataset.

Clasificador	Parámetros	Accuracy	Accuracy balanceado	Sensibilidad	Especificidad	AUC
JRip	Promedio	0.4927	0.6265	0.4619	0.7912	0.4917
	Desv Est	0.0522	0.0459	0.0750	0.0208	0.1257
	Mejor	0.5854	0.7009	0.5855	0.8325	0.7222
	Peor	0.3902	0.5426	0.3202	0.7410	0.2778
OneR	Promedio	0.4577	0.5982	0.4171	0.7792	0.5894
	Desv Est	0.0692	0.0649	0.1051	0.0271	0.1417
	Mejor	0.5610	0.6939	0.5669	0.8209	0.875
	Peor	0.2195	0.4335	0.1645	0.7026	0.3056
PART	Promedio	0.4122	0.5877	0.4015	0.7739	0.5556
	Desv Est	0.0514	0.0502	0.0817	0.0214	0.1292
	Mejor	0.5366	0.7073	0.5899	0.8248	0.7778
	Peor	0.2927	0.4923	0.2336	0.7351	0.3194

promedios con la aplicación de los filtros superaron en su mayoría (9 de 12 modelos) al resultado anterior. El promedio más alto fue de 0.6564.

En la Figura 1 se observan los mejores resultados que pertenecen a los clasificadores con todas las variables (+) y a la aplicación de los filtros seleccionados. Se obtuvo el valor más alto de 0.6564 con el clasificador OneR y con los filtros Chi.squared (▲) y RandomForest (⊠) que aparece con la abreviatura RanFor en la imagen. La línea punteada intercepta el número 0.6564 en el eje y, que representa al valor más alto alcanzado por los modelos mencionados anteriormente. Se observa en el eje x, el clasificador OneR y sobre él, los puntos más altos alcanzado por los filtros Chi.squared y RandomForest, superpuestos uno sobre el otro. Esto debido a que los dos lograron el mismo resultado promedio.

Del mismo modo se aprecia que debajo del mayor resultado alcanzado con todas las variables (+) solo se encuentran los modelos JRip con filtro Consistency (■) y PART con filtros CFS (●) y Consistency (■) por lo que se puede definir que mediante la selección de atributos por los métodos filtros se incrementa el rendimiento con respecto a los modelos con todas las variables.

5. Conclusión

El objetivo de este trabajo fue observar el comportamiento de los clasificadores basados en reglas para la creación de un modelo de diagnóstico efectivo para el SGB con variables clínicas. Para ello aplicamos tres clasificadores basados en reglas (JRip, OneR, PART) y cuatro métodos de selección de atributos relevantes (Chi.squared, CFS, Consistency, RandomForest). Realizamos experimentos usando todas las variables del conjunto de datos y usando únicamente las variables relevantes identificadas por cada filtro.

Para este estudio comparamos el rendimiento de los modelos elaborados en este trabajo, con el parámetro del *accuracy* balanceado. El mejor rendimiento

Tabla 4. Rendimiento de los modelos utilizando las variables relevantes identificadas por cada filtro.

Chi.squared						
Clasificador	Parámetros	Accuracy	Accuracy balanceado	Sensibilidad	Especificidad	AUC
JRip	Promedio	0.5398	0.6549	0.5023	0.8075	0.5097
	Desv Est	0.0554	0.0453	0.0717	0.0218	0.1047
	Mejor	0.6585	0.7336	0.6272	0.8501	0.7153
	Peor	0.4146	0.5179	0.2785	0.7573	0.2500
OneR	Promedio	0.5122	0.6564	0.5113	0.8014	0.6588
	Desv Est	0.0675	0.0518	0.0795	0.0255	0.1317
	Mejor	0.6098	0.7294	0.6272	0.8402	0.8611
	Peor	0.3659	0.5544	0.3509	0.7503	0.3056
PART	Promedio	0.5065	0.6561	0.5143	0.7979	0.5736
	Desv Est	0.0646	0.0549	0.0857	0.0251	0.1187
	Mejor	0.6098	0.7599	0.6711	0.8488	0.8333
	Peor	0.3415	0.5219	0.2982	0.7455	0.3889
CFS						
Clasificador	Parámetros	Accuracy	Accuracy balanceado	Sensibilidad	Especificidad	AUC
JRip	Promedio	0.5293	0.6463	0.4899	0.8027	0.5100
	Desv Est	0.0521	0.0523	0.0845	0.0222	0.1138
	Mejor	0.6341	0.7371	0.6272	0.847	0.7083
	Peor	0.439	0.5179	0.2785	0.7525	0.2778
OneR	Promedio	0.4894	0.6319	0.4712	0.7926	0.616
	Desv Est	0.0617	0.0479	0.0756	0.0225	0.1549
	Mejor	0.6098	0.7294	0.6195	0.8402	0.875
	Peor	0.3659	0.544	0.3224	0.7503	0.3056
PART	Promedio	0.4561	0.6241	0.4615	0.7868	0.5475
	Desv Est	0.0651	0.0499	0.0795	0.0227	0.1357
	Mejor	0.561	0.705	0.5844	0.8277	0.8125
	Peor	0.2927	0.526	0.3114	0.7363	0.3472
Consistency						
Clasificador	Parámetros	Accuracy	Accuracy balanceado	Sensibilidad	Especificidad	AUC
JRip	Promedio	0.4813	0.6098	0.4394	0.7803	0.5352
	Desv Est	0.0693	0.0436	0.0673	0.0221	0.1096
	Mejor	0.6098	0.7142	0.5987	0.8298	0.7431
	Peor	0.2683	0.4557	0.2018	0.7097	0.2500
OneR	Promedio	0.5041	0.6473	0.4974	0.7972	0.647
	Desv Est	0.0728	0.0625	0.0965	0.0298	0.1334
	Mejor	0.6098	0.7294	0.6272	0.8402	0.8611
	Peor	0.3415	0.4508	0.1919	0.7096	0.3056
PART	Promedio	0.4333	0.5997	0.4264	0.773	0.5435
	Desv Est	0.0743	0.0422	0.0632	0.0249	0.0815
	Mejor	0.5610	0.6932	0.5669	0.8196	0.7083
	Peor	0.2927	0.5146	0.3059	0.7233	0.3611
RandomForest						
Clasificador	Parámetros	Accuracy	Accuracy balanceado	Sensibilidad	Especificidad	AUC
JRip	Promedio	0.5260	0.6540	0.5083	0.7996	0.5495
	Desv Est	0.0499	0.0408	0.0643	0.0193	0.0906
	Mejor	0.6829	0.7684	0.6743	0.8624	0.7500
	Peor	0.4390	0.5739	0.3695	0.7660	0.3333
OneR	Promedio	0.5122	0.6564	0.5113	0.8014	0.6588
	Desv Est	0.0675	0.0518	0.0795	0.0255	0.1317
	Mejor	0.6098	0.7294	0.6272	0.8402	0.8611
	Peor	0.3659	0.5544	0.3509	0.7503	0.3056
PART	Promedio	0.4992	0.6483	0.501	0.7956	0.5595
	Desv Est	0.0683	0.0445	0.067	0.0239	0.1257
	Mejor	0.5854	0.7299	0.6272	0.8416	0.8611
	Peor	0.3415	0.5317	0.3268	0.7366	0.2778

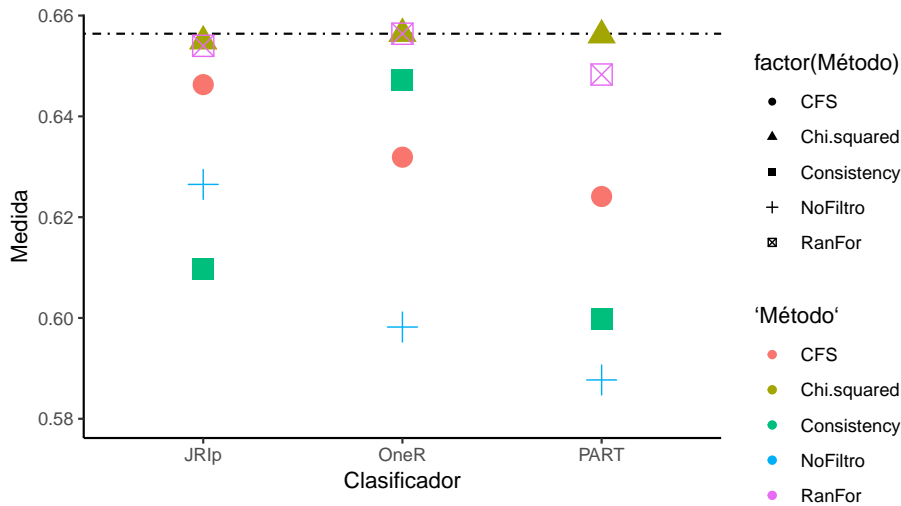


Fig. 1. Comparación de los promedios del *accuracy* balanceado de cada modelo.

obtenido fue de 0.6564 con el clasificador OneR y los filtros Chi.squared y RandomForest.

Es de mucho interés nuestro el poder construir un clasificador con un *accuracy* balanceado con un porcentaje de confiabilidad mejorado, para asegurar el diagnóstico de los 4 subtipos de SGB. Por lo tanto se propone continuar con diferentes experimentos para mejorar este resultado como algoritmos de árbol de decisión, de regresión, selección de atributos mediante *wrappers*, y demás, para apoyar de mejor manera a esta área de la medicina.

Agradecimientos Al CONACYT por el apoyo a la Maestría en Ciencias de la Computación de la Universidad Juárez Autónoma de Tabasco.

Referencias

1. Canul-Reich, J., Frausto-Solís, J., Hernández-Torruco, J.: A predictive model for Guillain-Barré syndrome based on single learning algorithms. *Computational and Mathematical Methods in Medicine* 2017, 1–9 (2017)
2. Canul-Reich, J., Hernández-Torruco, J., Frausto-Solís, J., Méndez Castillo, J.J.: Finding relevant features for identifying subtypes of Guillain-Barré syndrome using quenching simulated annealing and partitions around medoids. *International Journal of Combinatorial Optimization Problems and Informatics* 6(2), 11–27 (2015)
3. Cohen, W.: Fast effective rule induction. In: *Proceedings of the Twelfth International Conference on Machine Learning*. pp. 115–123. Morgan Kaufmann (1995)
4. Cruz-Montegudo, M., Cordeiro, M.N.D., Borges, F.: Computational chemistry approach for the early detection of drug-induced idiosyncratic liver toxicity. *Journal of computational chemistry* 29(4), 533–549 (2008)

5. Farid, D.M., Al-Mamun, M.A., Manderick, B., Nowe, A.: An adaptive rule-based classifier for mining big biological data. *Expert Systems with Applications* 64, 305–316 (2016)
6. Frank, E., Witten, I.H.: Generating accurate rule sets without global optimization. In: *Proceedings of the Fifteenth International Conference*. pp. 144–151. University of Waikato, Department of Computer Science (1998)
7. Gu, J., Ghosal, S.: Bayesian ROC curve estimation under binormality using a rank likelihood. *Journal of statistical planning and inference* 139(6), 2076–2083 (2009)
8. Han, J., Pei, J., Kamber, M.: *Data mining: concepts and techniques*. Elsevier (2011)
9. Hernández-Torruco, J., Canul-Reich, J., Frausto-Solís, J., Méndez-Castillo, J.J.: Towards a predictive model for Guillain-Barré syndrome. In: *Proceedings of the 37th Annual International Conference of the IEEE Engineering in Medicine and Biology Society*. pp. 7234–7237. Milano, Italy (2015)
10. Hindle, A., German, D.M., Godfrey, M.W., Holt, R.C.: Automatic classification of large changes into maintenance categories. In: *IEEE 17th International Conference on Program Comprehension*. pp. 30–39 (2009)
11. Holte, R.C.: Very simple classification rules perform well on most commonly used datasets. *Machine Learning* 11(1), 63–90 (1993)
12. Hornik, K., Buchta, C., Hothorn, T., Karatzoglou, A., Meyer, D., Zeileis, A., Hornik, M.K.: *Package RWeka* (2020)
13. Instituto Mexicano del Seguro Social: Diagnóstico y tratamiento: Síndrome de Guillain-Barré, segundo y tercer nivel de atención. *Catálogo Maestro de Guías de Práctica Clínica*. IMSS (2016)
14. Jiao, Y., Du, P.: Performance measures in evaluating Machine Learning based bioinformatics predictors for classifications. *Quantitative Biology* 4(4), 320–330 (2016)
15. Keleş, M.K., Kılıç, Ü.: Artificial Bee Colony algorithm for feature selection on SCADI dataset. In: *3rd International Conference on Computer Science and Engineering (UBMK)*. pp. 463–466. IEEE (2018)
16. Kuhn, M., Wing, J., Weston, S., Williams, A., Keefer, C., Engelhardt, A., Cooper, T., Mayer, Z., Kenkel, B., Team, R.C., et al.: *Package caret*. *The R Journal* (2020)
17. Peric, S., Milosevic, V., Berisavac, I., Stojiljkovic, O., Beslac-Bumbasirevic, L., Marjanovic, I., Djuric, V., Djordjevic, G., Rajic, S., Cvijanovic, M.: Clinical and epidemiological features of Guillain-Barré syndrome in the Western Balkans. *Journal of the peripheral nervous system* 19(4), 317–321 (2014)
18. Rendell, L., Seshu, R.: Learning hard concepts through constructive induction: Framework and rationale. *Computational Intelligence* 6(4), 247–270 (1990)
19. Revelle, W., Revelle, M.W.: *Package psych*. *The Comprehensive R Archive Network* (2015)
20. Robin, X., Turck, N., Hainard, A., Tiberti, N., Lisacek, F., Sanchez, J.C., Müller, M., Siegert, S., Doering, M., Robin, M.X.: *Package pROC*. *Tech. rep.* (2020)
21. Romanski P., Kotthoff L., .K.M.L.: *Selecting Attributes Package Fselector*. CRAN (2018)
22. Sokolova, M., Lapalme, G.: A systematic analysis of performance measures for classification tasks. *Information processing & management* 45(4), 427–437 (2009)
23. Steinberg, J.S., Koski, C.L.: Síndrome de Guillain-Barré, polineuropatía desmielinizante inflamatoria aguda (PDIA) y sus variantes. *The GBS/CIDP Foundation International* (2010)
24. Torres-Vásquez, M., Chávez-Bosquez, O., Hernández-Ocaña, B., Hernández-Torruco, J.: Classification of guillain-barré syndrome subtypes using sampling techniques with binary approach. *Symmetry* 12(3), 482 (2020)

25. Torres-Vásquez, M., Hernández-Torruco, J., Hernández-Ocana, B., Chávez-Bosquez, O.: Balanceo de datos del síndrome de Guillain-Barré utilizando SMOTE para la clasificación de subtipos. *Research in Computing Science* 148(7), 113–125 (2019)
26. Urbanek, S., Urbanek, M.S., JDK, S.J.: Package *rJava* (2019)
27. Witten, I., Frank, E., Hall, M.A.: *Data Mining: Practical Machine Learning Tools and techniques*. Morgan Kaufmann, third edn. (2011)
28. World Health Organization: Identification and treatment of Guillain-Barré syndrome in the context of the Zika virus outbreak. WHO 1 (2016)