Análisis de métodos de clasificación del sueño usando KNN y regresión lineal

Bourlot, Jimena; Eberle, Gerónimo; Garay, Ramiro

*Procesamiento Digital de Señales, Ingeniería en Informática, Facultad de Ingeniería y Ciencias Hídricas, Universidad Nacional del Litoral.*

***Resumen. El objetivo de este proyecto fue obtener una comparación entre un método basado en regresión lineal, y un método KNN, que permiten clasificar el estado de una persona en “despierto” o “dormido” de acuerdo a distintas variables de interés.***

***En primer lugar reprodujimos el método de la regresión lineal a partir de los datos de la aceleración de una persona, que fue estudiado por Cole y otros. Para ello, hicimos un preprocesamiento de la señal, luego un ventaneo y obtención de características de interés. Posteriormente, en base a un entrenamiento del método con un grupo de prueba y comparación con el resultado de etiquetas del sueño obtenidas en base a un estudio de PSG, pudimos probarlo con un grupo distinto de personas y así poder analizar la especificidad, sensibilidad, y tasa de aciertos del método.***

***Una vez realizado esto, nos dedicamos a implementar un método de K Nearest Neighbours (KNN), el cual realizaba el mismo procedimiento de clasificar el estado de una persona en “despierto” o “dormido” en base a distintas variables. Para este caso, realizamos nuevamente un entrenamiento con un grupo de entrenamiento para ajustar el método, y luego con un grupo de prueba pudimos analizar su eficiencia.***

***Una vez realizado todo esto, comparamos el método de la regresión lineal usando aceleración con el método KNN con la misma variable. Luego, como añadido, comparamos también el uso de esta variable con el de otras de forma individual, y finalmente comparamos los resultados con una combinación entre las variables nombradas anteriormente, de forma tal que pudimos analizar si el método mejora al añadir determinadas variables o si la adición no tiene efectos notables.***

***Palabras clave: sueño - HR - aceleración - regresión - KNN.***

1. INTRODUCCIÓN

El objetivo de este trabajo es realizar un análisis de dos métodos investigados, de forma tal que podamos compararlos en cuanto a su eficiencia en la clasificación de dos estados de una persona al dormir: despierto-dormido.

El primero de estos métodos es uno basado en regresión lineal de 7 variables relacionadas con la aceleración del movimiento de una persona al dormir [1].

El segundo de ellos es un método de K nearest neighbours (KNN), en el cual utilizaremos la aceleración de la misma forma que en el primer método, con el objetivo de compararlos; y luego añadiremos otras variables para analizar el funcionamiento del método con las mismas de manera individual; y luego utilizando una combinación de ellas.

El resultado final que esperamos obtener es saber si el método de la aceleración produce buenos resultados en comparación a un método convencional KNN que sí los tiene; y luego observar qué sucede al añadir otras variables, por ejemplo, si los resultados producen un efecto notorio; o si la aceleración es una buena medida para clasificar estados.

1. DATOS UTILIZADOS

Los datos fueron recolectados por la Universidad de Michigan desde junio de 2017 a marzo de 2019 y estuvieron involucrados 31 sujetos.

Los participantes fueron provistos de un Apple Watch con el cual dormirían mientras un polisomnógrafo recababa datos del sueño a la par. El polisomnógrafo es el método de “oro” para la clasificación del sueño, pero su utilización podría alterar las condiciones normales del paciente, y requiere la presencia de personal calificado.

Se quitaron del experimento aquellas personas que demostraban tener algún inconveniente al dormir a causa del PSG, ya que les produciría desorden en las etapas de REM. También se descartaron aquellas que poseyeran apnea obstructiva.

A partir de los datos de PSG, se tienen etiquetadas las etapas del sueño en la que se halla cada persona mediante un número del -1 al 5 que se corresponden con:

* -1: ninguna fase;
* 0: despierto;
* 1: etapa N1;
* 2: etapa N2;
* 3: etapa N3;
* 5: REM;

También el dataset cuenta con los datos de la aceleración obtenidos del Apple Watch, representados por las componentes (aceleración en dirección *x*, *y* y *z*), y un cuarto valor que corresponde al tiempo en segundos desde el 1º de enero de 1970. La aceleración en cada dirección está expresada en unidades de g.

Los datos de aceleración se encuentran muestreados a 50 Hz, mientras que los de ritmo cardíaco a 0,2 Hz aproximadamente, aunque con notables huecos de información. Esto último explicará la baja cobertura de los predictores que utilizan datos del ritmo cardíaco.

1. PREPROCESAMIENTO DE LOS DATOS

Realizamos un preprocesamiento de los datos que consistió en eliminar del dataset utilizado aquellas etiquetas, aceleración o pulso cardíaco que correspondiera a tiempos negativos de los datos del polisomnógrafo, dado que en esos tiempos el aparato aún no había comenzado a funcionar y no se producía un etiquetado que nos permitiera obtener resultados en esos tiempos. Lo mismo se realizó para tiempos posteriores a que el PSG dejara de etiquetar los estados.

Además, convertimos las etiquetas del polisomnógrafo, que originalmente indicaría una de 5 opciones, para que ahora sea del tipo “despierto” o “dormido”.

Adicionalmente, realizamos un procesamiento en las medidas de la aceleración, de forma tal que en vez de tener datos de la aceleración en cada dirección, tendremos la magnitud de la aceleración en cada instante de tiempo.

1. VENTANEO Y OBTENCIÓN DE CARACTERÍSTICAS

Se utilizaron ventanas de 30 segundos de tiempo para todas las variables utilizadas, menos para la entropía, que se utilizaron ventanas de 60. El motivo de esta elección es que la base de datos empleada tenía los estados de sueño categorizados por PSG en ventanas de 30, pero para el cálculo de entropía, que utilizaba datos del ritmo cardíaco muestreados cada 5 segundos, una ventana de tal largo albergaba un número muy pequeño de mediciones (aproximadamente 6) por ventana.

Debido a la forma en que los relojes recolectan los datos, cada ventana de tiempo tiene distinto número de mediciones, con lo cual debimos utilizar el tiempo de cada medición para asignarla a una ventana y hacer los cálculos correspondientes. Aquellas ventanas donde no se pudo obtener el valor de una característica, se puso el valor *-3*.

Para una mejor organización de los datos, y debido a que nos será útil al momento de aplicar los métodos, generamos una matriz de datos del paciente, que en cada fila tiene la etiqueta de PSG, junto con todas las características extraídas en esa ventana.

Las características extraídas fueron:

* Actividad del acelerómetro (máximo de cada ventana de tiempo).
* Actividad del ritmo cardíaco (máximo de cada ventana de tiempo).
* Entropía aproximada del ritmo cardíaco (calculado en ventanas de 60).

Las dos actividades fueron normalizadas por la media aritmética de todas las mediciones de su respectivo paciente.

La última característica fue extraída con la esperanza de utilizarla como un marcador de “confiabilidad” de los datos del ritmo cardíaco, cuyos resultados fueron muy pobres en los predictores KNN implementados.

1. MÉTODO DE LA REGRESIÓN

Para este método, estudiado por Roger J. Cole y otros [1], nos propusimos reproducirlo de acuerdo a los parámetros que se utilizó en aquella investigación. La ecuación de aquel método surgía de una regresión lineal de 7 variables de la siguiente forma:

<math xmlns="http://www.w3.org/1998/Math/MathML"><mi>D</mi><mo>&#xA0;</mo><mo>=</mo><mo>&#xA0;</mo><mn>0</mn><mo>.</mo><mn>0001</mn><mo>&#xA0;</mo><mo>(</mo><mn>404</mn><mo>&#xD7;</mo><msub><mi>A</mi><mrow><mo>-</mo><mn>4</mn></mrow></msub><mo>+</mo><mn>598</mn><msub><mi>A</mi><mrow><mo>-</mo><mn>3</mn></mrow></msub><mo>&#xA0;</mo><mo>+</mo><mn>326</mn><msub><mi>A</mi><mrow><mo>-</mo><mn>2</mn></mrow></msub><mo>+</mo><mn>441</mn><msub><mi>A</mi><mrow><mo>-</mo><mn>1</mn></mrow></msub><mo>+</mo><mn>1</mn><mo>.</mo><mn>408</mn><mo>&#xD7;</mo><msub><mi>A</mi><mn>0</mn></msub><mo>+</mo><mo>&#xA0;</mo><mn>508</mn><msub><mi>A</mi><mrow><mo>+</mo><mn>1</mn></mrow></msub><mo>+</mo><mn>350</mn><msub><mi>A</mi><mrow><mo>+</mo><mn>2</mn></mrow></msub><mo>)</mo></math>

<math xmlns="http://www.w3.org/1998/Math/MathML"><mi>D</mi><mo>&#xA0;</mo><mo>=</mo><mo>&#xA0;</mo><mn>0</mn><mo>.</mo><mn>0001</mn><mo>&#xA0;</mo><mo>(</mo><mn>404</mn><mo>&#xD7;</mo><msub><mi>A</mi><mrow><mo>-</mo><mn>4</mn></mrow></msub><mo>+</mo><mn>598</mn><msub><mi>A</mi><mrow><mo>-</mo><mn>3</mn></mrow></msub><mo>&#xA0;</mo><mo>+</mo><mn>326</mn><msub><mi>A</mi><mrow><mo>-</mo><mn>2</mn></mrow></msub><mo>+</mo><mn>441</mn><msub><mi>A</mi><mrow><mo>-</mo><mn>1</mn></mrow></msub><mo>+</mo><mn>1</mn><mo>.</mo><mn>408</mn><mo>&#xD7;</mo><msub><mi>A</mi><mn>0</mn></msub><mo>+</mo><mo>&#xA0;</mo><mn>508</mn><msub><mi>A</mi><mrow><mo>+</mo><mn>1</mn></mrow></msub><mo>+</mo><mn>350</mn><msub><mi>A</mi><mrow><mo>+</mo><mn>2</mn></mrow></msub><mo>)</mo></math>

la variable D representa el puntaje en cada ventana de tiempo. Y para calcularla se utilizan los valores de la ventana actual, 4 valores anteriores y dos siguientes.

Debido a que el método obtenía valores entre 0 y 2, y nuestro objetivo era clasificar de acuerdo a “despierto” (1) y “dormido” (0); realizamos el ajuste de esta variable de acuerdo a un umbral que nos permitía identificar a partir de qué valores podría considerarse despierto al paciente en estudio. El resultado fue graficado para tener una visión de cómo funciona el método (figura 1).

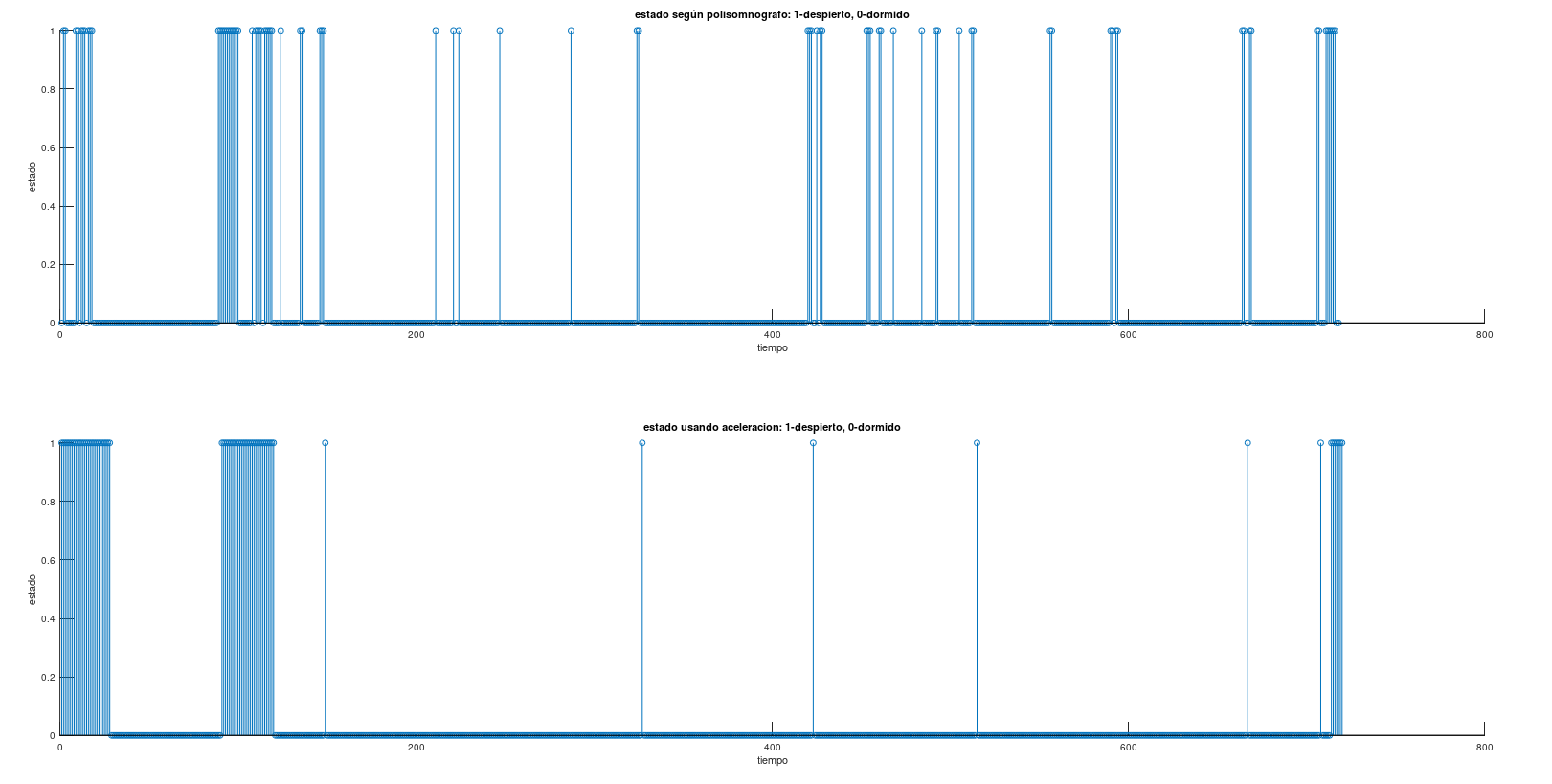


Figura 1. Funcionamiento del método usando aceleración en comparación con las etiquetas de PSG normalizadas a un puntaje 0-1 (“despierto” o “dormido”).

Para la elección del umbral, utilizamos el percentil que optimiza el valor del estadístico *F1* (que relaciona la especificidad y sensibilidad) para un grupo de pacientes de entrenamiento. Con este fin, usamos la función de minimización local de octave fminbnd(), calculando un error de la forma:

1. KNN

El método KNN (*K Nearest Neighbours)* consiste en tomar las características extraídas de un grupo de pacientes (grupo de entrenamiento) a lo largo de todos sus periodos de sueño, y utilizarlas como vectores en un espacio. De esta forma, dado un vector de características correspondientes a la medición de un paciente en un instante de tiempo, es posible buscar el vector más cercano (distancia euclídea) en el espacio de los vectores del grupo de entrenamiento. El parámetro *k* del método especifica el número de vecinos (vectores) más cercanos a considerar. La etiqueta más representada entre los vecinos (dormido o despierto) es la asignada a la medición en cuestión.

Debido a la posibilidad de empates si se elige un número par de vecinos, el valor de *k* utilizado fue siempre impar, eliminando este problema.

Estos vectores están conformados por una o más características de las detalladas en la sección IV, cada una de estas combinaciones originando un clasificador KNN distinto (mirar Tabla 1 para ver los distintos clasificadores).

1. ENTRENAMIENTO Y PRUEBA DE LOS MÉTODOS

Tanto el método de Cole y otros como el de KNN propuesto en este texto requieren de ciertos parámetros que deben ser obtenidos del dataset: en el primero se necesita definir un valor umbral para el clasificador, mientras que en el segundo es necesario construir matrices que contengan los vectores sobre los cuales computar las distancias en KNN.

Para comparar justamente ambos métodos, se usó el mismo grupo de entrenamiento y prueba para los dos.

En un principio se ejecutó el método KNN utilizando sólo la aceleración como variable de decisión, para poder comparar sus resultados con los del método de Cole y otros.

Posteriormente, se probó el método KNN para las variables adicionadas de forma individual. Finalmente se ejecutó para una combinación de estas variables, de forma tal que pudiéramos comparar los efectos en el resultado del método, al añadir cada una de estas variables.

La tabla que se puede encontrar al final de este texto contiene los resultados de todos los clasificadores promediados sobre 10 pruebas.

Para el entrenamiento de los métodos separamos el dataset disponible de 31 pacientes en distintos grupos de entrenamiento y prueba.

Para ambos métodos, se realizaron 10 experimentos de entrenamiento-prueba, en donde el grupo de entrenamiento es de 15 personas elegidas aleatoriamente, y el grupo de prueba es de las 16 personas restantes en cada caso.

Para el caso de la regresión, el entrenamiento fue realizado para hallar el percentil que nos dé el mayor valor de estadístico F1.

Para el caso del método KNN, los grupos de entrenamiento fueron empleados para el cálculo de los vectores que posteriormente nos permitirían decidir el estado del paciente en cada ventana de tiempo.

Las estadísticas finales de cada método son un promedio del resultado de cada experimento de prueba.

Para cada método se calcularon las siguientes medidas:

* Especificidad;
* Sensibilidad;
* Accuracy (exactitud);
* F1;
* Cobertura: este fue un indicador del porcentaje de ventanas en las que se contaba con datos para calcular el estado del paciente.

Nos basamos principalmente en los valores de F1 (que relaciona la sensibilidad con la especificidad) para sacar conclusiones al respecto de qué tan buenos son los resultados de cada método.

1. CONCLUSIONES

Como muestran los resultados finales, el método de la regresión lineal se desempeña considerablemente mejor que cualquiera de los clasificadores implementados por KNN, los últimos apenas acercándose a la mitad de la sensibilidad. Sin embargo, para los clasificadores que usan solo aceleración, los valores de especificidad son mayores y el valor de F1 (que sopesa tanto la sensibilidad y especificidad) no son mucho menores. Además, su exactitud es comparable a la del método de regresión y también tienen una cobertura mayor.

El HR no es una buena característica y se desempeña pobremente sola, y en conjunto con la aceleración sólo empeora los resultados obtenidos. En general, parece no haber una correlación muy grande entre ritmo cardíaco y el estadío de sueño del paciente.

Mientras que para valores mayores de k, los predictores basados únicamente en aceleración se desempeñaban mejor, para los basados en HR lo opuesto es cierto. Esto parece confirmar lo dicho previamente.

La adición de la entropía aproximada a los clasificadores basados en ritmo cardíaco no afectó notablemente los resultados obtenidos, apenas mejorando un poco el estadístico F1. Esto hizo descartar la posibilidad de un clasificador que involucre las tres características a la vez.

Por último, es de notar que hacia mayores valores de k, los predictores basados únicamente en aceleración aumentan de sensibilidad sin comprometer la especificidad obtenida (nótese el aumento de F1). Una mejora al trabajo actual presentado sería analizar el comportamiento de estos predictores para rangos de k aún mayores.

Una posible explicación del pobre desempeño de los clasificadores KNN es la pérdida de la información temporal. Mientras que el método de la regresión lineal considera la sucesión de valores en el tiempo, los clasificadores por KNN no.

Un resumen de estos resultados puede verse en la tabla 1.

|  |  |  |  |  |  |  |
| --- | --- | --- | --- | --- | --- | --- |
| K | Método de predicción | Sensibilidad | Especificidad | Cobertura | Accuracy | F1 |
| 0 | Paper | 0.546 | 0.936 | 0.963 | 0.891 | 0.469 |
| 9 | aceleración | 0.247 | 0.979 | 0.984 | 0.897 | 0.322 |
| 9 | Ritmo Cardiaco | 0.120 | 0.908 | 0.648 | 0.822 | 0.105 |
| 9 | aceleración+Ritmo Cardiaco | 0.234 | 0.970 | 0.637 | 0.888 | 0.267 |
| 9 | Ritmo Cardíaco+ApEn (r = 3) | 0.113 | 0.926 | 0.166 | 0.783 | 0.113 |
| 9 | Ritmo Cardíaco+ApEn (r = 4) | 0.112 | 0.919 | 0.166 | 0.777 | 0.107 |
| 9 | Ritmo Cardíaco+ApEn (r = 5) | 0.116 | 0.912 | 0.166 | 0.772 | 0.100 |
| 9 | Ritmo Cardíaco+ApEn (r = 6) | 0.116 | 0.912 | 0.166 | 0.772 | 0.100 |
| 11 | aceleración | 0.249 | 0.979 | 0.984 | 0.898 | 0.325 |
| 11 | Ritmo Cardiaco | 0.093 | 0.940 | 0.648 | 0.848 | 0.096 |
| 11 | aceleración+Ritmo Cardiaco | 0.235 | 0.975 | 0.637 | 0.892 | 0.275 |
| 11 | Ritmo Cardíaco+ApEn (r = 3) | 0.099 | 0.935 | 0.166 | 0.788 | 0.100 |
| 11 | Ritmo Cardíaco+ApEn (r = 4) | 0.103 | 0.926 | 0.166 | 0.781 | 0.098 |
| 11 | Ritmo Cardíaco+ApEn (r = 5) | 0.105 | 0.921 | 0.166 | 0.777 | 0.090 |
| 11 | Ritmo Cardíaco+ApEn (r = 6) | 0.105 | 0.921 | 0.166 | 0.777 | 0.090 |
| 13 | aceleración | 0.250 | 0.979 | 0.984 | 0.898 | 0.327 |
| 13 | Ritmo Cardiaco | 0.078 | 0.947 | 0.648 | 0.853 | 0.083 |
| 13 | aceleración+Ritmo Cardiaco | 0.233 | 0.976 | 0.637 | 0.894 | 0.274 |
| 13 | Ritmo Cardíaco+ApEn (r = 3) | 0.087 | 0.949 | 0.166 | 0.799 | 0.097 |
| 13 | Ritmo Cardíaco+ApEn (r = 4) | 0.089 | 0.931 | 0.166 | 0.784 | 0.092 |
| 13 | Ritmo Cardíaco+ApEn (r = 5) | 0.084 | 0.935 | 0.166 | 0.787 | 0.080 |
| 13 | Ritmo Cardíaco+ApEn (r = 6) | 0.084 | 0.935 | 0.166 | 0.787 | 0.080 |
| 15 | aceleración | 0.252 | 0.980 | 0.984 | 0.899 | 0.330 |
| 15 | Ritmo Cardiaco | 0.079 | 0.950 | 0.648 | 0.856 | 0.084 |
| 15 | aceleración+Ritmo Cardiaco | 0.230 | 0.980 | 0.637 | 0.897 | 0.281 |
| 15 | Ritmo Cardíaco+ApEn (r = 3) | 0.084 | 0.950 | 0.166 | 0.799 | 0.095 |
| 15 | Ritmo Cardíaco+ApEn (r = 4) | 0.078 | 0.952 | 0.166 | 0.799 | 0.086 |
| 15 | Ritmo Cardíaco+ApEn (r = 5) | 0.091 | 0.939 | 0.166 | 0.791 | 0.091 |
| 15 | Ritmo Cardíaco+ApEn (r = 6) | 0.091 | 0.939 | 0.166 | 0.791 | 0.091 |
| 17 | aceleración | 0.252 | 0.980 | 0.984 | 0.899 | 0.330 |
| 17 | Ritmo Cardiaco | 0.068 | 0.964 | 0.648 | 0.867 | 0.077 |
| 17 | aceleración+Ritmo Cardiaco | 0.230 | 0.982 | 0.637 | 0.898 | 0.285 |
| 17 | Ritmo Cardíaco+ApEn (r = 3) | 0.074 | 0.959 | 0.166 | 0.805 | 0.085 |
| 17 | Ritmo Cardíaco+ApEn (r = 4) | 0.073 | 0.955 | 0.166 | 0.802 | 0.083 |
| 17 | Ritmo Cardíaco+ApEn (r = 5) | 0.087 | 0.946 | 0.166 | 0.795 | 0.087 |
| 17 | Ritmo Cardíaco+ApEn (r = 6) | 0.087 | 0.946 | 0.166 | 0.795 | 0.087 |
| 21 | aceleración | 0.255 | 0.980 | 0.984 | 0.899 | 0.333 |
| 21 | Ritmo Cardiaco | 0.060 | 0.979 | 0.648 | 0.880 | 0.071 |
| 21 | aceleración+Ritmo Cardiaco | 0.224 | 0.982 | 0.637 | 0.898 | 0.279 |
| 21 | Ritmo Cardíaco+ApEn (r = 3) | 0.055 | 0.969 | 0.166 | 0.811 | 0.065 |
| 21 | Ritmo Cardíaco+ApEn (r = 4) | 0.063 | 0.968 | 0.166 | 0.812 | 0.071 |
| 21 | Ritmo Cardíaco+ApEn (r = 5) | 0.072 | 0.956 | 0.166 | 0.802 | 0.074 |
| 21 | Ritmo Cardíaco+ApEn (r = 6) | 0.072 | 0.956 | 0.166 | 0.802 | 0.074 |

Tabla 1. Comparación de los métodos utilizados de acuerdo a su especificidad, sensibilidad, accuracy y cobertura.

1. REFERENCIAS

[1] Roger J. Cole, Daniel F. Kripke, William Gruen, Daniel J. Mullaney and J. Christian Gillin, "Automatic Sleep/Wake Identification From Wrist Activity". San Diego Veterans Affairs Medical Center, San Diego, California 92093, U.S.A.; and Ambulatory Monitoring, Inc., Ardsley, New York 10502, U.S.A. Sleep. Vol. 15. No.5. 1992.

[2]Ramiro Casala, Leandro E. Di Persia, Gastón Schlotthauera, “Sleep-wake stages classification using heart rate signals from pulse oximetry”. Lab. de Señales y Dinámicas no Lineales, Facultad de Ingeniería, Universidad Nacional de Entre Ríos (UNER), Argentina. Consejo Nacional de Investigaciones Científicas y Técnicas (CONICET), Argentina. Instituto de Investigación en Señales, Sistemas e Inteligencia Computacional - Universidad Nacional del Litoral - CONICET. Instituto de Investigación y Desarrollo en Bioingeniería y Bioinformática - UNER - CONICET. sinc(i) Research Institute for Signals, Systems and Computational Intelligence (fich.unl.edu.ar/sinc). Heliyon, Vol 5, No 10, pp e02529, 2019.

[3] Walch, O. (2019) 'Motion and heart rate from a wrist-worn wearable and labeled sleep from polysomnography' (version 1.0.0), PhysioNet. Available at: <https://doi.org/10.13026/hmhs-py35>.