

---

# Reconocimiento de la actividad física a través de sensores

Alicia Brown Rivero-Borrell  
Regina Ceballos Mondragón

*“Numbers have an important story to tell.  
They rely on you to give them a clear and convincing voice.”  
- Stephen Few*

## Introducción

En el mundo más de 5 mil millones de personas tienen un smartphone y se prevé que este número incremente en el futuro (según la *GSMA*). Los teléfonos inteligentes pasan en nuestra mano, bolsa o cerca de nosotros, prácticamente todo el día y, debido a los sensores de movimiento, posición y orientación que tienen incorporados, reciben señales de la actividad (o ausencia de) que realizamos; con los datos recopilados se puede generar mucha información acerca de nuestros patrones de conducta.

El presente estudio pretende desarrollar un modelo que reconozca la actividad que realiza un individuo cualquiera, en un momento específico, usando la base de datos de UCI, *Smartphone-Based Recognition of Human Activities and Postural Transitions Data Set* [1]. En particular, queremos clasificar con precisión si un individuo está caminando, subiendo o bajando escaleras, sentado, parado o acostado.

Los resultados de un análisis como éste son tienen muchos beneficios y aplicaciones, principalmente dirigido hacia desarrollo de un sistema de salud inteligente [10]. Ejemplos de ellos son las aplicaciones y dispositivos de monitoreo de la salud, rendimiento o monitoreo de un usuario mediante la medición precisa de su actividad física. A pesar de la comodidad que representan los teléfonos celulares, éstos no han superado a los wearable devices [7] en la precisión.

## Descripción de los datos

La base [1] comprende 10929 observaciones y 561 atributos, resultados del experimento realizado en 30 voluntarios, de entre 19 y 48 años, a quienes se les pidió hacer: tres posturas estáticas (de pie, sentado, acostado), tres actividades dinámicas (caminar, bajar escaleras y subir escaleras) y transiciones posturales entre dichas posturas (pararse-a-sentarse, pararse-a-recostarse, sentarse-a-pararse, sentarse-a-recostarse, recostarse-a-pararse y recostarse-a-sentarse). Todos los participantes llevaban un teléfono inteligente (*Samsung Galaxy S II*) en la cintura durante la ejecución del experimento.

A pesar de tener sólo 30 individuos, la cantidad de información es extensa porque se le hicieron varias transformaciones, adiciones y modificaciones a la original. Primero, se capturó la aceleración lineal de 3 ejes y la velocidad angular de 3 ejes a una velocidad constante

de 50 Hz utilizando el acelerómetro integrado y el giroscopio del dispositivo (**tAcc-XYZ** y **tGyro-XYZ**<sup>1</sup>). La señal de aceleración se separó en señales de aceleración de cuerpo y gravedad (**tBodyAcc-XYZ** y **tGravityAcc-XYZ**). Posteriormente, la aceleración lineal del cuerpo y la velocidad angular se obtuvieron a tiempo para obtener señales de Jerk (**tBodyAccJerk-XYZ** y **tBodyGyroJerk-XYZ**<sup>2</sup>). También se calculó la magnitud de estas señales tridimensionales utilizando la norma euclidiana (**tBodyAccMag**, **tGravityAccMag**, **tBodyAccJerkMag**, **tBodyGyroMag**, **tBodyGyroJerkMag**). Después, se se aplicó una Transformada Rápida de Fourier (FFT) a algunas de estas señales produciendo **fBodyAcc-XYZ**, **fBodyAccJerk-XYZ**, **fBodyGyro-XYZ**, **fBodyAccJerkMag**, **fBodyGyroMag**, **fBodyGyroJerkMag**. Además, se estimaron: valor medio, desviación estándar, desviación media absoluta, mayor valor en matriz, valor más pequeño en matriz, área de magnitud de la señal, medida de energía<sup>3</sup>, rango intercuartílico, entropía de la señal, coeficientes de autorregresión, coeficiente de correlación entre dos señales, índice del componente de frecuencia con mayor magnitud, promedio ponderado de los componentes de frecuencia para obtener una frecuencia media, asimetría de la señal de dominio de frecuencia, curtosis de la señal de dominio de frecuencia, energía de un intervalo de frecuencia dentro de los 64 intervalos de la FFT de cada ventana y ángulo entre vectores. Por último, promediando las señales en una muestra de ventana de señal se obtuvieron los vectores: **GravityMean**, **tBodyAccMean**, **tBodyAccJerkMean**, **tBodyGyroMean**, **tBodyGyroJerkMean**<sup>4</sup>.

Por otro lado, para cada observación se tiene la etiqueta de la actividad en la que se tomó. Éstas son:

1 – WALKING	2 – WALKING UPSTAIRS	3 – WALKING DOWNSTAIRS
4 – SITTING	5 – STANDING	6 – LAYING
7 – STAND TO SIT	8 – SIT TO STAND	9 – SIT TO LIE
10 – LIE TO SIT	11 – STAND TO LIE	12 – LIE TO STAND

## Metodología

1. **Limpieza:** no se encontraron observaciones con variables incompletas, atípicas o vacías, por lo tanto ninguna observación se eliminó o modificó.
2. **Integración:** la base de datos estaba previamente dividida en un conjunto de entrenamiento (70 %) y un conjunto de prueba (30 %). A nuestra percepción, esto le quitaba aleatoriedad a los conjuntos y decidimos integrarlos. Posteriormente, hacemos una selección de conjuntos aleatoria, como se menciona a continuación.
3. **Selección:** para crear el conjunto de entrenamiento, validación y de prueba, seleccionamos aleatoriamente a las personas, en vez de las observaciones. Esto debido a que el propósito de nuestro modelo es poder predecir qué actividad está realizando alguien,

<sup>1</sup>El prefijo t es para indicar el tiempo, medida continua.

<sup>2</sup>“-XYZ” se utiliza para denotar señales de 3 ejes en las direcciones X, Y y Z.

<sup>3</sup>Suma de los cuadrados divididos por el número de valores.

<sup>4</sup>Las unidades utilizadas para las aceleraciones (total y cuerpo) son g (gravedad de la Tierra =  $9.80665m/seg^2$ ) y las unidades del giroscopio son  $rad/seg$ .

---

independientemente de la forma en que se mueve, así que consideramos que era necesario que si una persona estaba en el conjunto de entrenamiento, debía estar toda la información que le corresponde y así eliminar el posible sesgo que provenga de la manera particular de moverse.

El número de observaciones varía entre persona, entonces cada vez que se hace el ejercicio de selección el tamaño de los conjuntos cambia.

La estructura de los conjuntos es la siguiente:

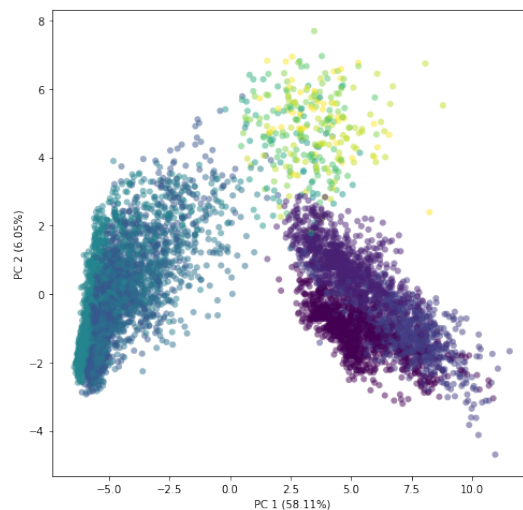
- Entrenamiento: 18 personas (alrededor del 65 %)
- Validación: 5 personas (alrededor del 15 %)
- Prueba: 7 personas (alrededor del 20 %)

4. **Transformación:** las variables están normalizadas y limitadas dentro de  $[-1, 1]$ , por lo que tampoco transformamos los datos. Igualmente, al no saber la importancia de cada variable, decidimos mantener todos los atributos para comenzar el análisis.
5. **Minado:**[2][3][4]

### PCA

Para poder ver la estructura y comportamiento de los datos, necesitamos reducir la dimensión. Usamos PCA para encontrar las 2 componentes principales. Encontramos que la primera explica alrededor del 57 % de la varianza total y la segunda 6.65 %, por lo que en conjunto tenemos más de la mitad de la variación explicada. Transformando los datos a ese nuevo sistema de ejes, conseguimos la siguiente gráfica.

Figura 1: Datos sobre las 2 componentes principales



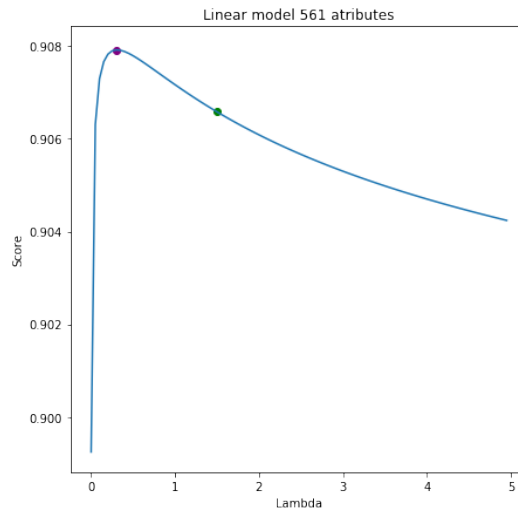
---

Los colores corresponden a la etiqueta del tipo de actividad que se realiza, por lo que, en **Figura 1**, podemos ver una división entre los grupos de actividades de movimiento (azules), actividades estáticas (moradas) y transiciones (amarillas-verdes). No obstante, algunas actividades se superponen por completo o de manera parcialmente. Por lo que necesitamos maximizar los márgenes al separar estas actividades para dar un resultado más eficiente.

### **Linear model**

Por la manera en que se modificó la base de datos y por lo que indica el resultado del PCA, creemos que hay atributos que no son relevantes o que no agregan información extra al modelo, por lo que es posible eliminar algunos y reducir así la complejidad del problema. Para esto, hacemos una regresión lineal y validamos para encontrar el parámetro lambda.

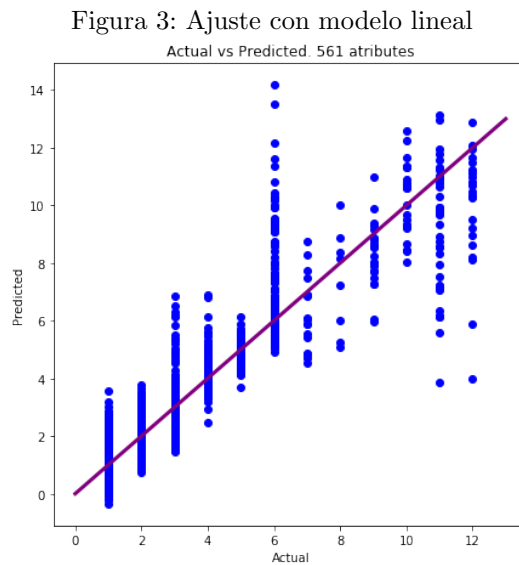
Figura 2: Score al cambiar lambda



En **Figura 2** notamos que con  $\text{lambda} = 0.3$  (marcado en la figura con un punto morado), el *score* es el máximo por lo que procedemos a entrenar los datos con *regularización de Ridge*. Una vez hecho esto, nos quedamos con aquellos coeficientes que en valor absoluto sean mayores a 1. Con lo que logramos reducir de 561 atributos a 63. Sin embargo, notamos que podemos tomar un  $\text{lambda}$  mayor, por ejemplo  $\text{lambda} = 1.5$  (marcado en la figura con un punto verde) para reducir el número de atributos más. Con esto logramos reducir los atributos relevantes a 25 y sólo perdimos centésimas en la precisión por lo que tomamos este como nuestro modelo.

---

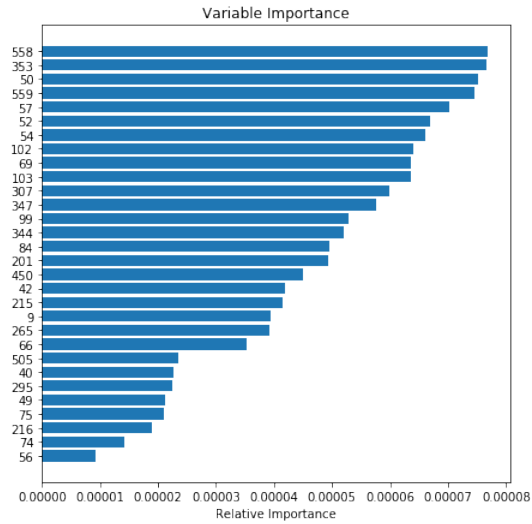
En **Figura 3** se grafica el valor real contra el valor estimado después de la regresión. A través de la gráfica, se percibe que el modelo ajusta mejor a las categorías 1-6 de actividades dinámicas y posturas estáticas (los puntos son más cercanos a la recta identidad) que a las categorías de transición.



### **Random Forest**

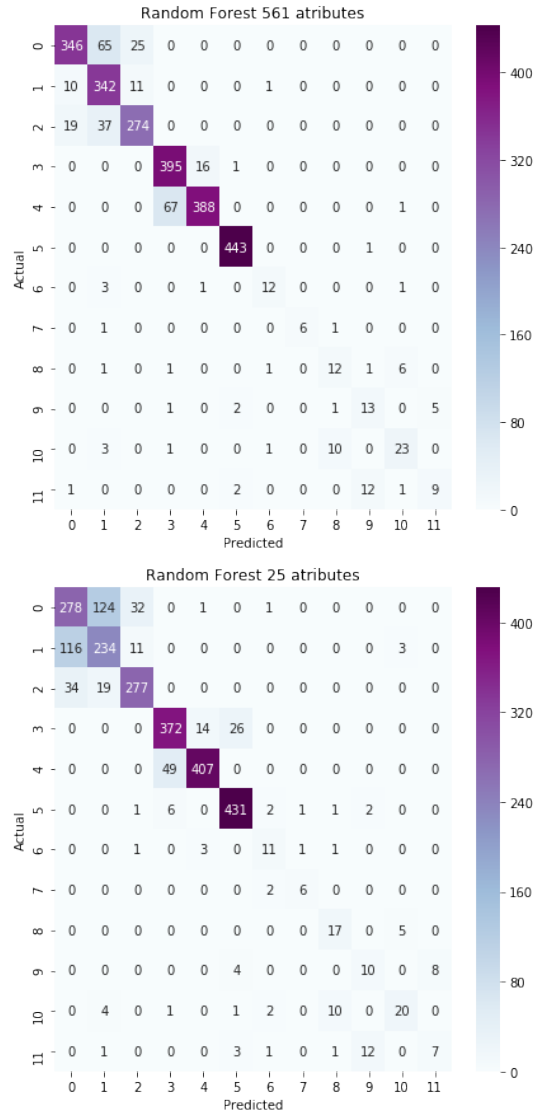
Siguiendo la línea de reducción de dimensión, procedemos a hacer el análisis con Random Forest. Ordenamos a las variables según su importancia y vemos que, a partir de la 29 tienen menos del .001 % de importancia (Véase **Figura 4**)

Figura 4: Importancia



Además, utilizamos Random Forest para clasificar los datos del conjunto de prueba. En **Figura 5** se muestra la diferencia entre el mapa de calor utilizando todos los atributos y utilizando sólo aquéllos que obtuvimos como relevantes en el modelo lineal.

Figura 5: Comparación entre heatmaps de Random Forest

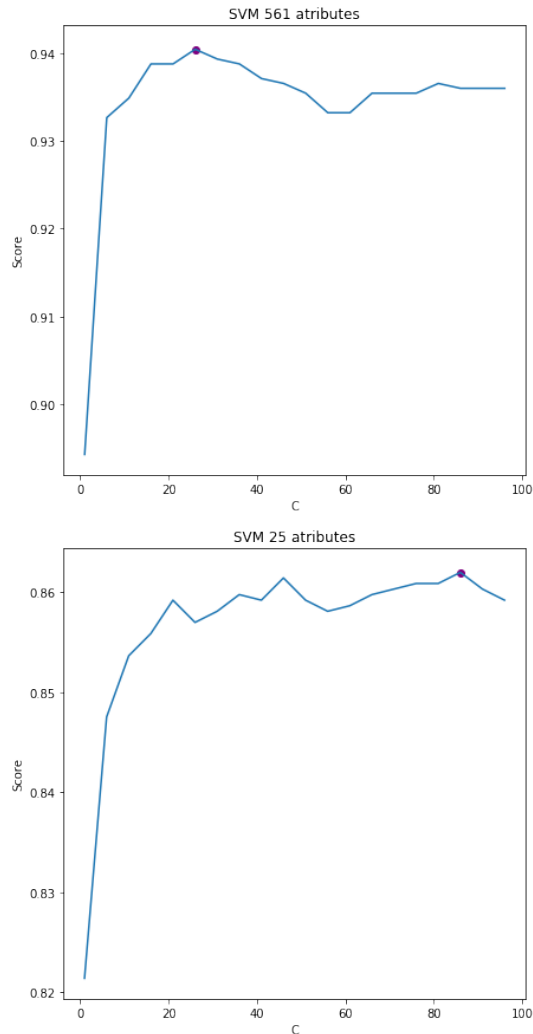


Se puede ver que se alcanza un gran nivel de precisión, pues la mayor parte de los datos fue clasificado correctamente. La categoría 6 (*laying*) es la que alcanzó casi el 100% de aciertos, mientras que la 1 (*walking*) es la que más errores de clasificación tiene. La eliminación de más de 500 variables hace que la precisión sea un poco menor pero no cambia radicalmente, es decir, se mantiene dentro de un rango *acceptable*.

## SVM

Para separar el las categorías que los datos parecen inducir, utilizamos SVM. Primero hicimos validación para encontrar la  $C$  que minimiza el error y ajustamos con SVM los datos (completos y reducidos). Después, comparamos el desempeño de los Kernels: “linear”, “RBF”, “polynomial” (grado 3) (véase Figura 6, Figura 7<sup>5</sup>).

Figura 6: Comparación entre score SVM



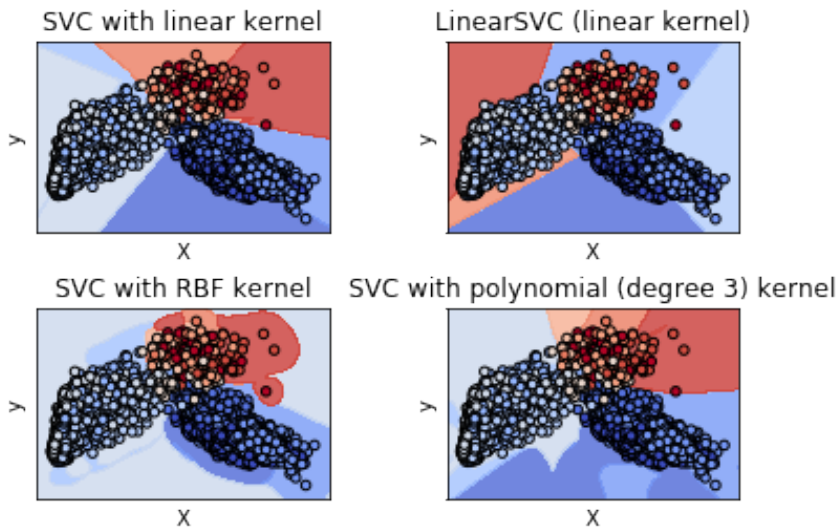
<sup>5</sup>Código de [http://scikit-learn.org/stable/auto\\_examples/svm/plot\\_iris.html](http://scikit-learn.org/stable/auto_examples/svm/plot_iris.html)



---

El movimiento curva del Score cambia al disminuir los atributos, sin embargo se sigue alcanzando un *score* aceptable con menos atributos. Cabe notar que la *C* aumenta al disminuir los atributos. Una explicación a este fenómeno es que el modelo busca compensar la falta de información a través de mayor penalización de errores.

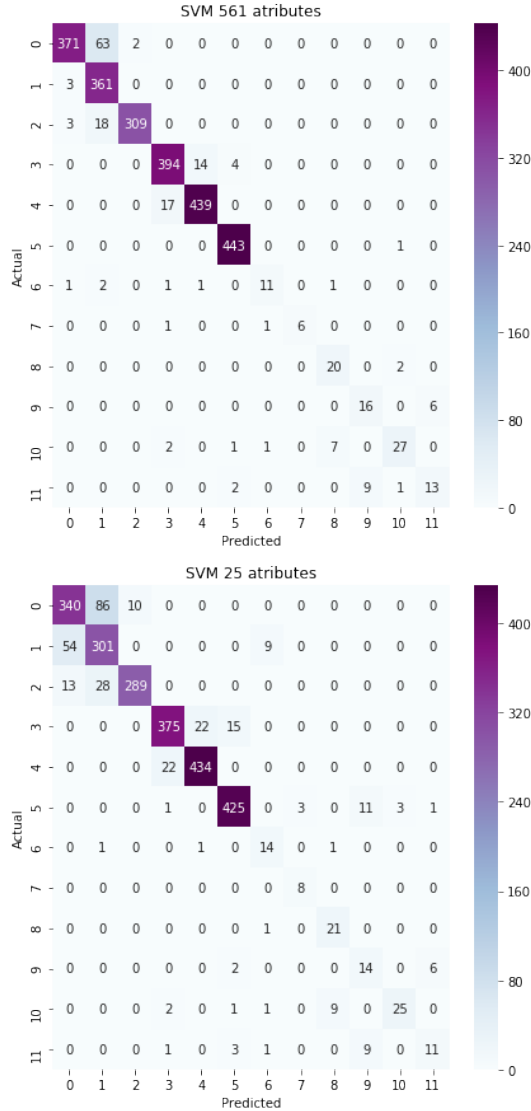
Figura 7: Comparación Kernels



Como era de esperarse, el Kernel que hace la división de *clusters* más precisa es RBF (Gaussiano), ya que el espacio del kernel tiene dimensión infinita.

En la **Figura 8** se muestra el cambio en los mapas de calor de SVM y análogamente a lo sucedido en Random Forest, se reduce la precisión pero no significativamente.

Figura 8: Comparación entre heatmaps de SVM



### K Nearest Neighbors:

Con la intención de encontrar la técnica que mejor ajusta y generaliza los datos, ahora hacemos K Nearest Neighbors, comenzando por la validación para encontrar la cantidad de vecinos que debemos usar. Posteriormente, ajustamos y generamos graficas de score (Figura 9) y mapas de calor.(Figura 10)

Figura 9: Comparación entre score KNN

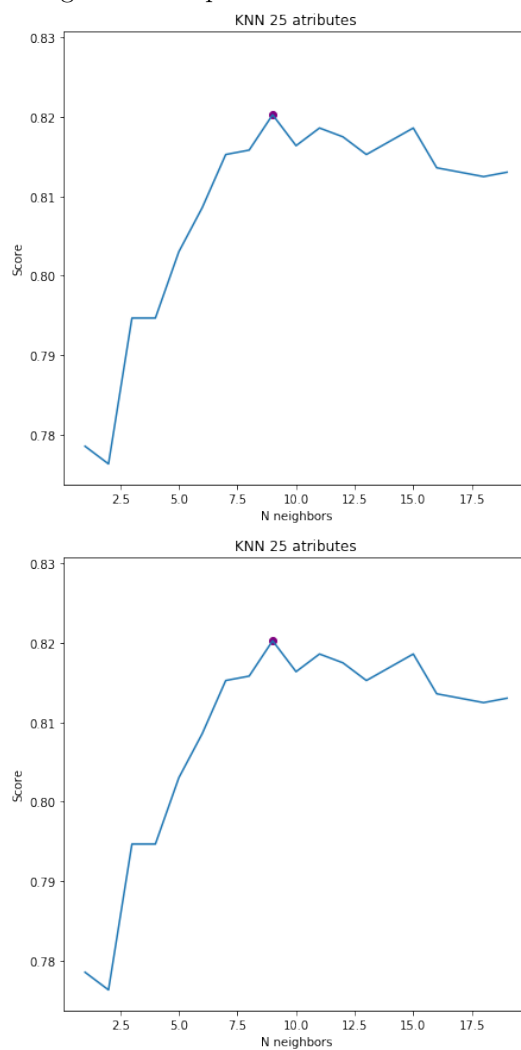
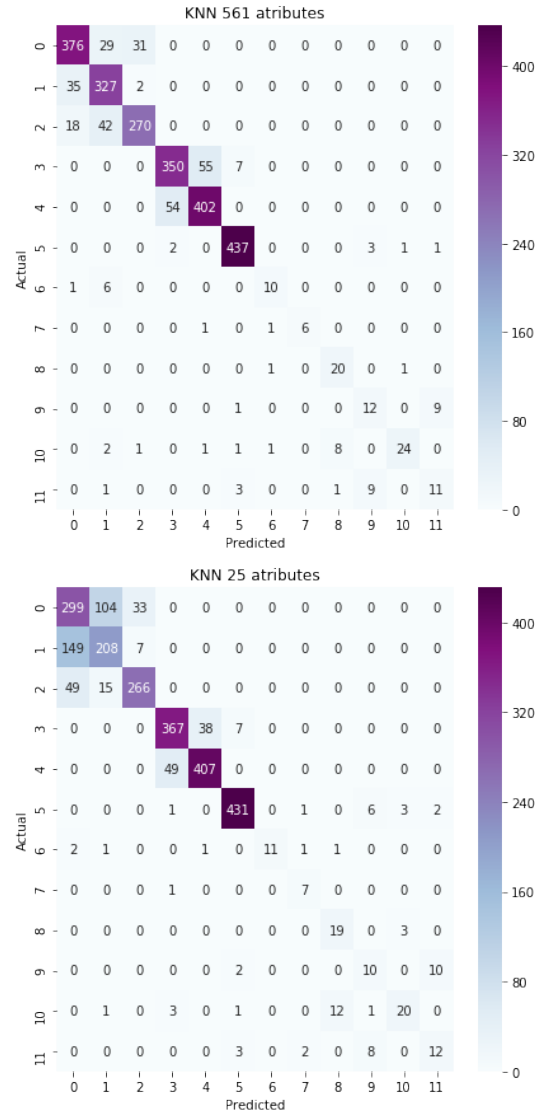


Figura 10: Comparación entre heatmap KNN



Obtenemos resultados similares a los de los modelos anteriores.

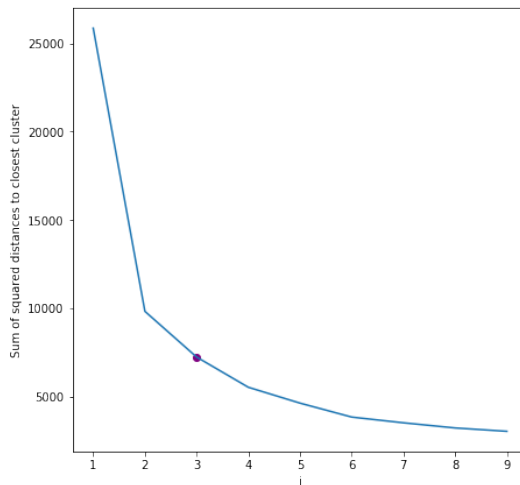
### K Means

Por último, utilizamos K-Means para intentar mejorar el desempeño de RF, SVM y/o KNN. Hacemos validación para hallar el número de medias a estimar (Figura 11) y

---

luego aplicamos el algoritmo a los modelos.

Figura 11: Dist K-Means



En esta ocasión, la implementación de K-Means empeora el resultado, consideramos que esto proviene de la similitud entre grupos de datos.

## 6. Resumen de resultados:

Nuestro análisis buscaba maximizar la precisión de las predicciones y la siguiente tabla muestra las estadísticas obtenidas para cada uno de las técnicas utilizadas para facilitar su comparación y elegir aquella con mejores resultados.<sup>6</sup>

---

<sup>6</sup>Recordemos que  $Score = (Precision * Recall) / (2Precision + Recall)$ , donde:  
 $Precision = \text{correct predictions of class} / \text{total predictions of class}$  y  
 $Recall = \text{correct predictions of class} / \text{total observations of class}$ .

MODELO	No. ATRIBUTOS	PARÁMETROS	SCORE TRAIN	SCORE TEST
Random Forest	561	–	99.93 %	87.92 %
Random Forest	25	–	99.81 %	80.42 %
Random Forest + KMeans	25	k=3	46.39 %	43.82 %
KNN	561	N=5	97.76 %	87.22 %
KNN	25	N=9	92.77 %	79.91 %
KNN + KMeans	25	k=3 N=9	35.62 %	32.56 %
SVM	561	C=26	98.58 %	93.63 %
SVM	25	C=86	92.65 %	87.68 %
SVM + KMeans	25	k=3 C=86	33.81 %	35.24 %

## Conclusión

Basandonos en los *scores*, podemos ver que la clasificación es extremadamente precisa independientemente del modelo si se utilizan todas las variables y no se utiliza KMeans. Un resultado sorprendente es que el uso de KMeans deteriora inmensamente la precisión. A esto se le puede atribuir, como se mencionó anteriormente, la similaridad entre grupos de actividades.

Por otro lado, consideramos que la disminución de variables a través de la regresión Ridge fue satisfactoria, ya que el *score* en todos los modelos se mantiene arriba del 80 %. En general, creemos que el propósito del minado se logró ya que las clasificaciones son en su mayoría correctas. Al tener el mayor *score*, pensamos que el modelo SVM con reducción de atributos se puede implementar en aplicaciones de la vida real. Sin embargo, creemos que sería oportuno mejorar la precisión en el caso de que la aplicación sea de salud.

La alta precisión que encontramos en las clasificaciones aunado con el aumento de celulares, sugiere un incremento en aplicaciones de clasificación de actividad humana. Algunos de los problemas posibles a los que se puede enfrentar la aplicación de modelos de este estilo para la clasificación de la actividad humana son los siguientes:

- El usuario necesariamente debe utilizar el celular para poder analizar su actividad. Es por esto que los wearable devices tiene mayor precisión.
- Se debe buscar ampliar el análisis a uno en que las actividades no sean secuenciales. Es decir, un análisis en el que se tome en cuenta la realización de múltiples actividades a la vez.
- La repetición de una actividad por ejemplo sentarse y pararse muchas veces puede ser en realidad otra actividad: hacer sentadillas.

- La interrupción de actividades puede llevar a confusiones. Por ejemplo: ir caminando y caerse se puede ver como caminar y recostarse.
- El uso del celular por múltiples personas.
- El mal uso de información confidencial.
- Una clasificación errónea puede llevar a un mal diagnóstico en temas de salud.

[6]

## Referencias

- [1] UCI-Smartphone-Based Recognition of Human Activities and Postural Transitions Data Set  
<http://archive.ics.uci.edu/ml/datasets/Smartphone-Based+Recognition+of+Human+Activities+and+Postural+Transitions#>  
 (última consulta en diciembre de 2017).
- [2] Christopher M. Bishop. Pattern Recognition and Machine Learning. Springer. 2006.  
<http://users.isr.ist.utl.pt/~wurmd/Livros/school/Bishop%20-%20Pattern%20Recognition%20And%20Machine%20Learning%20-%20Springer%20%202006.pdf>  
 (última consulta en diciembre de 2017).
- [3] Trevor Hastie, Robert Tibshirani, Jerome Friedman. The Elements of Statistical Learning Data Mining, Inference, and Prediction. Springer. 2d edition. <https://web.stanford.edu/~hastie/Papers/ESLII.pdf>  
 (última consulta en diciembre de 2017).
- [4] Fernando Esponda. Presentaciones el repositorio ClaseML2017. <https://github.com/ClaseML-2017/MaterialyTareas/tree/master/clases>  
 (última consulta en diciembre de 2017).
- [5] Davide Anguita<sup>1</sup>, Alessandro Ghio<sup>1</sup>, Luca Oneto<sup>1</sup>, Xavier Parra<sup>2</sup> and Jorge L. Reyes-Ortiz<sup>1,2</sup>. A Public Domain Dataset for Human Activity Recognition Using Smartphones. ESANN 2013 proceedings, European Symposium on Artificial Neural Networks, Computational Intelligence and Machine Learning. Bruges (Belgium), 24-26 April 2013. <https://www.elen.ucl.ac.be/Proceedings/esann/esannpdf/es2013-84.pdf>  
 (última consulta en diciembre de 2017).
- [6] Eunju Kim, Sumi Helal and Diane Cook. Human Activity Recognition and Pattern Discovery  
<http://citeseerx.ist.psu.edu/viewdoc/download?doi=10.1.1.154.6429&rep=rep1&type=pdf>  
 (última consulta en diciembre de 2017).

- 
- [7] Brent Rose. No, Phones Aren't More Accurate Than Fitness Wearables. 03.18.1512:00 PM.  
<https://www.wired.com/2015/03/fitness-tracking-test/>  
(última consulta en diciembre de 2017).
- [8] Jessica Moorea and Binghai Lingb, Human Activity Recognition using Smartphone Sensors. Final Project. December 2016.  
<http://cs229.stanford.edu/proj2016/report/LingMoore-HumanActivityRecognitionUsingS.pdf>  
(última consulta en diciembre de 2017).
- [9] Nicholas Canova and Fjoralba Shemaj. Human Activity Recognition using Smartphone Sensor Data. Final Project. December 2016.  
<http://cs229.stanford.edu/proj2016/report/CanovaShemaj-HumanActivityRecognitionUsi.pdf>  
(última consulta en diciembre de 2017).
- [10] Mohd Fikri Azli bin Abdullah, Ali Fahmi Perwira Negara, Md. Shohel Sayeed, Deok-Jai Choi, Kalaifarasi Sonai Muthu. Classification Algorithms in Human Activity Recognition using Smartphones. World Academy of Science, Engineering and Technology. International Journal of Biomedical and Biological. Engineering. Vol:6, No:8, 2012.  
<http://www.waset.org/publications/8520>  
(última consulta en diciembre de 2017).