

Reconocimiento de Actividades Humanas Utilizando Teléfonos Inteligentes

Salgado López Álvaro¹

UANL - Facultad De Ciencias Físico Matemáticas
Av. Universidad s/n. Ciudad Universitaria San Nicolás de los Garza - México

1. Introducción

El Reconocimiento de Actividades Humanas (HAR, por sus siglas en inglés) constituye un esfuerzo fundamental en el ámbito de comprender el comportamiento humano a través de medios tecnológicos. Su objetivo fundamental radica en identificar las acciones realizadas por individuos basándose en una colección de observaciones derivadas tanto del individuo como de su entorno circundante. Tradicionalmente, las metodologías de reconocimiento han aprovechado diversas fuentes de datos, que van desde señales ambientales hasta sensores colocados en el cuerpo, para lograr este objetivo. Destacadamente, los sensores de movimiento especializados ubicados en varias partes del cuerpo, como la cintura, la muñeca, el pecho y los muslos, han demostrado un rendimiento de clasificación notable. Sin embargo, estas configuraciones de sensores convencionales a menudo conllevan incomodidad para los usuarios y no ofrecen una solución sostenible a largo plazo para el monitoreo de actividades.

El advenimiento de los teléfonos inteligentes ha inaugurado una nueva era de oportunidades de investigación en aplicaciones centradas en el ser humano. Con los usuarios como fuentes de información contextual y los teléfonos inteligentes como herramientas de detección primarias, estos dispositivos ubicuos equipados con una variedad de sensores integrados, incluidos acelerómetros, giroscopios, micrófonos y cámaras duales, presentan una alternativa atractiva para el HAR. Al aprovechar los sensores inerciales incorporados en los teléfonos inteligentes, los investigadores han comenzado a explorar vías para monitorear automáticamente y de manera discreta las Actividades de la Vida Diaria (AVD). Notablemente, los teléfonos inteligentes brindan una solución flexible, asequible y autónoma que integra perfectamente el monitoreo de actividades con servicios de telefonía [1].

En este trabajo, se exploran técnicas de aprendizaje no supervisado para clasificar diferentes tipos de actividades físicas a partir de datos recogidos por sensores. A diferencia de los métodos supervisados, los algoritmos no supervisados no requieren un conjunto de datos previamente etiquetados para entrenar el modelo, lo cual puede ser ventajoso en escenarios donde la obtención de datos etiquetados es costosa o impráctica.

2. Marco Teórico

En el campo de la clasificación de actividades humanas mediante datos de sensores, es crucial seleccionar y evaluar adecuadamente los modelos de aprendizaje automático. La clasificación de actividades humanas tiene aplicaciones en diversos dominios como la salud, el deporte, la seguridad y la interacción humano-computadora. Los datos de sensores, como acelerómetros y giroscopios, proporcionan información importante sobre los movimientos y posturas de una persona, lo que permite identificar diferentes actividades.

2.1. Métricas de Evaluación

Para evaluar el rendimiento de los modelos de clasificación, es esencial utilizar métricas adecuadas que proporcionen una visión completa del desempeño del modelo. Las métricas más comunes en el contexto de la clasificación de actividades humanas son:

Exactitud

La exactitud mide la proporción de predicciones correctas sobre el total de predicciones realizadas.

$$\text{Exactitud} = \frac{\text{Número de predicciones correctas}}{\text{Número total de predicciones}} \quad (1)$$

Precisión

La precisión mide la proporción de verdaderos positivos sobre el total de predicciones positivas.

$$\text{Precisión} = \frac{\text{Verdaderos Positivos}}{\text{Verdaderos Positivos} + \text{Falsos Positivos}} \quad (2)$$

Sensibilidad o Tasa de Verdaderos Positivos

La sensibilidad mide la proporción de verdaderos positivos sobre el total de positivos reales.

$$\text{Sensibilidad} = \frac{\text{Verdaderos Positivos}}{\text{Verdaderos Positivos} + \text{Falsos Negativos}} \quad (3)$$

F1-Score

El *F1-Score* es la media armónica de la precisión y la sensibilidad, proporcionando un balance entre ambos.

$$\text{F1-Score} = 2 \times \frac{\text{Precisión} \times \text{Sensibilidad}}{\text{Precisión} + \text{Sensibilidad}} \quad (4)$$

Matriz de Confusión

La matriz de confusión es una tabla que muestra la distribución de predicciones correctas e incorrectas, categorizadas por cada clase. Es una herramienta fundamental para analizar el rendimiento del modelo en cada clase individual.

3. Metodología

Se llevaron a cabo una serie de experimentos para obtener el conjunto de datos HAR. Se seleccionó un grupo de 30 voluntarios con edades comprendidas entre los 19 y 48 años para esta tarea. A cada persona se le indicó seguir un protocolo de actividades mientras llevaba un teléfono inteligente Samsung Galaxy S II montado en la cintura. Las seis actividades de la vida diaria (ADL) seleccionadas fueron estar de pie, sentarse, acostarse, caminar, bajar escaleras y subir escaleras. Cada sujeto realizó el protocolo dos veces: en el primer intento, el teléfono inteligente se colocó en el lado izquierdo del cinturón, y en el segundo, el usuario mismo lo colocó según su preferencia. También hubo una separación de 5 segundos entre cada tarea, donde se les indicaba a los individuos descansar, lo que facilitó la repetibilidad (cada actividad se intentó al menos dos veces) y la generación de datos de referencia a través de la interfaz visual.

No.	Estático	Tiempo (seg)	No.	Dinámico	Tiempo (seg)
0	Inicio (De pie Pos)	0	7	Caminar (1)	15
1	De pie (1)	15	8	Caminar (2)	15
2	Sentarse (1)	15	9	Bajar escaleras (1)	12
3	De pie (2)	15	10	Subir escaleras (2)	12
4	Acostarse (1)	15	11	Bajar escaleras (1)	12
5	Sentarse (2)	15	12	Subir escaleras (2)	12
6	Acostarse (2)	15	13	Bajar escaleras (3)	12
			14	Subir escaleras (3)	12
			15	Detenerse	0
Total					192

Cuadro 1: Protocolo de actividades para el Experimento HAR.

Las tareas se realizaron en condiciones de laboratorio, pero se les pidió a los voluntarios que realizaran libremente la secuencia de actividades para obtener un conjunto de datos más naturalista. El Cuadro 1 muestra los detalles del protocolo experimental [1].

3.1. Aprendizaje no supervisado

Además del popular algoritmo de aprendizaje no supervisado K-Means, una alternativa poderosa para abordar el reconocimiento de actividades humanas (HAR) en conjuntos de datos complejos es el modelo de mezclas gaussianas (GMM, por sus siglas en inglés).

El algoritmo de GMM es especialmente eficaz cuando se trabaja con datos que poseen características continuas, como aquellas obtenidas de sensores de

acelerómetro y giroscopio. A diferencia de K-Means, que asigna cada punto de datos a un solo clúster, GMM modela la distribución de los datos como una mezcla de varias distribuciones gaussianas. Esto significa que GMM puede capturar la complejidad de conjuntos de datos donde los puntos pueden pertenecer a múltiples clústeres o grupos.

Un GMM asume que los datos son generados a partir de una combinación (mezcla) de varias distribuciones gaussianas (normales), cada una con su propia media y varianza. Matemáticamente, esto se expresa como:

$$p(x) = \sum_{k=1}^K \pi_k \mathcal{N}(x|\mu_k, \Sigma_k) \quad (5)$$

Donde:

$\mathcal{N}(x|\mu_k, \Sigma_k)$ es una distribución gaussiana con media μ_k y covarianza Σ_k .

π_k es el peso de la k -ésima componente gaussiana (con $\sum_{k=1}^K \pi_k = 1$).

3.2. Aprendizaje supervisado

Se optó por utilizar el algoritmo de Máquinas de Vectores de Soporte (SVM, por sus siglas en inglés) debido a su eficacia en espacios de alta dimensión, lo cual es ideal para este conjunto de datos ya que contiene características extraídas de múltiples sensores de teléfonos inteligentes, como acelerómetros y giroscopios. Las características en este conjunto de datos pueden ser complejas y altamente dimensionales, y las SVM puede encontrar un hiperplano óptimo que separa eficazmente las diferentes actividades humanas en este espacio de características. Además, algunas variantes de las SVM han demostrado previamente su capacidad en problemas similares, obteniendo buenos resultados [2].

Las SVM utilizan una función de pérdida Hinge para maximizar el margen y minimizar el error de clasificación:

$$\min_{\mathbf{w}, b} \frac{1}{2} \|\mathbf{w}\|^2 + C \sum_{i=1}^n \max(0, 1 - y_i(\mathbf{w} \cdot \mathbf{x}_i + b)) \quad (6)$$

Donde:

$\|\mathbf{w}\|^2$ es la norma cuadrada del vector de pesos \mathbf{w} , C es el parámetro de regularización que controla el *trade-off* entre el margen y el error de clasificación, \mathbf{x}_i son las muestras de entrenamiento, y y_i son las etiquetas de clase correspondientes ($y_i \in \{-1, 1\}$).

4. Resultados

4.1. Aprendizaje no supervisado

Luego de implementar el algoritmo de GMM se obtuvo una clasificación de las actividades realizadas. Se empleó el algoritmo de GMM para abordar la

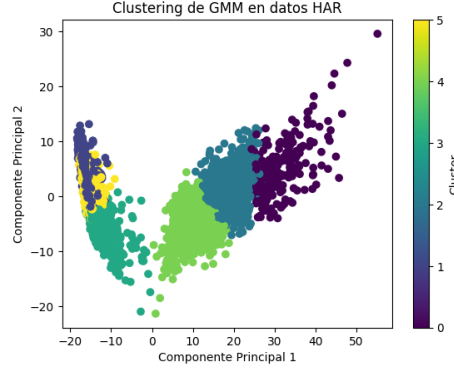


Figura 1: Agrupamiento de GMM en datos HAR

clasificación de las HAR. Los resultados de esta aplicación se presentan en la Figura 1.

Los hallazgos revelan una exitosa clasificación, donde se muestran claramente 6 grupos distintos en la Figura, cada uno representado con un color único. Cabe destacar que cada uno de estos grupos corresponde a una de las 6 actividades realizadas durante el estudio.

Este enfoque de clasificación basado en GMM permite una visualización clara y una interpretación precisa de los patrones de actividad identificados en los datos de HAR. La capacidad del algoritmo para modelar distribuciones complejas de datos y asignar probabilidades de pertenencia a múltiples agrupamientos ha demostrado ser efectiva en la identificación de diferentes actividades humanas, incluso en casos donde existen superposiciones o variaciones sutiles entre las actividades.

4.2. Aprendizaje Supervisado

Tras la implementación del algoritmo SVM, se logró alcanzar una precisión del 98 %. Este resultado refleja que el modelo pudo clasificar correctamente un alto porcentaje de los datos del conjunto de prueba, demostrando su eficacia en la tarea de reconocimiento de actividades humanas basadas en datos de sensores de teléfonos inteligentes. La alta precisión obtenida demuestra la capacidad del SVM para manejar de manera efectiva las características complejas y multidimensionales extraídas de los sensores, como acelerómetros y giroscopios. Además, la Figura 2 presenta detalladamente la matriz de confusión generada por el modelo. Esta matriz proporciona de manera visual cómo se clasificaron las diferentes actividades, mostrando el número de predicciones correctas e incorrectas para cada clase. Por ejemplo, se observa que la mayoría de las actividades fueron clasificadas con precisión casi perfecta.

El reporte de clasificación también destaca la sensibilidad (*recall*) del modelo de 98 %, que indica la capacidad del SVM para identificar correctamente las

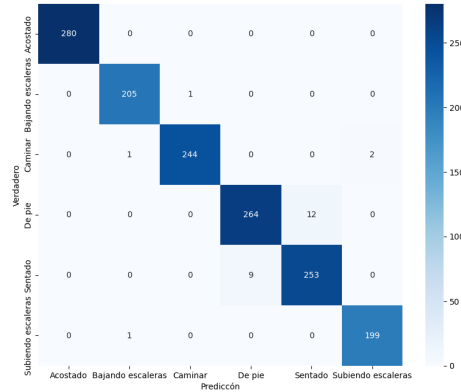


Figura 2: Matriz de Confusión

muestras de cada clase. En general, el modelo mostró una alta sensibilidad en todas las actividades, lo que refuerza su robustez y confiabilidad en aplicaciones prácticas.

En la literatura, los resultados del análisis se reportan típicamente en términos de precisión de clasificación utilizando varias métricas estándar como precisión, recall y F1-Score. En general, los estudios investigados reportan una precisión de clasificación muy alta, típicamente superior al 95 %. Por lo tanto, en el Cuadro 2 se presenta el reporte de clasificación obtenido, el cual proporciona una visión detallada del rendimiento del modelo para cada una de las clases en el problema de clasificación de actividades humanas.

Clase	Precisión	Recall	F1-Score	Soporte
De pie	1.00	1.00	1.00	280
Sentado	0.99	1.00	0.99	206
Acostado	1.00	0.99	0.99	247
Caminar	0.97	0.96	0.96	276
Bajando escaleras	0.95	0.97	0.96	262
Subiendo escaleras	0.99	0.99	0.99	200
Exactitud				0.98
Macro Avg	0.98	0.98	0.98	1471
Weighted Avg	0.98	0.98	0.98	1471

Cuadro 2: Reporte de Clasificación del Algoritmo SVM

El modelo demostró un rendimiento sobresaliente, con valores de precisión, *recall* y *F1-Score* consistentemente altos en todas las clases. Estos resultados indican que el modelo es altamente eficaz para la clasificación de actividades humanas, lo que lo hace idóneo para aplicaciones en entornos reales donde se requiere una identificación precisa y fiable de las actividades.

5. Conclusión

En este estudio, se exploró el uso del algoritmo de Gaussian Mixture Models (GMM) para clasificar las Actividades Humanas Reconocidas (HAR). El conjunto de datos utilizado en este estudio contiene el etiquetado de las actividades, lo que permitió una evaluación directa del desempeño del algoritmo sin necesidad de calcular el número de grupos.

Mediante el análisis de los resultados obtenidos, se ha demostrado la eficacia de GMM en la clasificación precisa de actividades humanas basadas en datos de sensores de acelerómetro y giroscopio. Los hallazgos revelaron que GMM es capaz de identificar patrones complejos en los datos de HAR y asignar de manera efectiva las observaciones a diferentes clústeres que representan actividades específicas. La capacidad de GMM para modelar la distribución de los datos de manera flexible y capturar las relaciones no lineales entre las características contribuyó significativamente a su rendimiento en la clasificación de actividades humanas.

Además, se observó que GMM proporciona una visualización clara de las actividades identificadas, lo que facilita la interpretación de los resultados y la identificación de patrones significativos. Esta capacidad de interpretación es crucial para aplicaciones prácticas en áreas como el monitoreo de la salud, la rehabilitación física y la optimización del rendimiento deportivo.

Adicionalmente, el SVM se presenta como una solución eficaz y confiable para el reconocimiento de actividades humanas, ofreciendo un alto nivel de precisión y capacidad para gestionar múltiples clases de actividades de manera efectiva. Estos resultados son prometedores para su implementación en sistemas de seguimiento y análisis de actividades diarias, proporcionando una solución robusta y precisa para el monitoreo basado en datos de sensores.

Referencias

- [1] V. Jacques, P. Filzmoser, C. Bernet, and C. Gosselin, "High-dimensional data clustering," in *Proceedings of the 21st European Symposium on Artificial Neural Networks, Computational Intelligence and Machine Learning (ESANN)*, 2013, pp. 345–350. [Online]. Available: <https://www.esann.org/sites/default/files/proceedings/legacy/es2013-84.pdf>
- [2] D. Anguita, A. Ghio, L. Oneto, X. Parra, and J. L. Reyes-Ortiz, "A public domain dataset for human activity recognition using smartphones," in *21st European Symposium on Artificial Neural Networks, Computational Intelligence and Machine Learning, ESANN 2013*, Bruges, Belgium, April 2013. [Online]. Available: https://link.springer.com/chapter/10.1007/978-3-030-51517-1_35