

Análisis, diseño, y simulación de un sistema de clasificación basado en técnicas de aprendizaje de máquina para Human Activity Recognition

Samuel Acevedo Bustamante
Ingeniería de Sistemas
Universidad de Antioquia
Medellín, Colombia
samuel.acevedob@udea.edu.co

Danilo Antonio Tovar Arias
Ingeniería de Sistemas
Universidad de Antioquia
Medellín, Colombia
danilo.tovar@udea.edu.co

Oswald Daniel Gutiérrez Cortina
Ingeniería de Sistemas
Universidad de Antioquia
Medellín, Colombia
odaniel.gutierrez@udea.edu.co

Resumen—Este documento contiene la información relacionada con el proceso de investigación, análisis y evaluación de diferentes modelos de inteligencia artificial para un conjunto de datos de Reconocimiento de Actividad Humana usando teléfonos inteligentes; a través de la exploración de artículos con diferentes aproximaciones para la solución de este problema y la realización de diferentes experimentos se busca comprender el proceso de modificación y utilización de modelos de inteligencia artificial.

Palabras Clave—Reconocimiento de Actividad Humana, Inteligencia Artificial, Clasificación

I. INTRODUCCIÓN

El reconocimiento automático de actividades humanas (En inglés, Human Activity Recognition, HAR) mediante sensores embebidos en dispositivos móviles ha emergido como un campo clave en el desarrollo de aplicaciones inteligentes para la salud, el deporte, la seguridad y la interacción humano-computador. En este contexto, el conjunto de datos *Human Activity Recognition Using Smartphones*, disponible públicamente en [1], representa una valiosa fuente de datos para la evaluación y comparación de algoritmos de aprendizaje automático en tareas de clasificación multiclase basadas en señales temporales.

Este dataset fue recopilado a partir de un experimento conducido con 30 voluntarios (19 hombres y 11 mujeres) de entre 19 y 48 años. Cada participante portó un *smartphone* "Samsung Galaxy S II" adherido a la cintura, mientras realizaba seis actividades físicas: caminar, subir escaleras, bajar escaleras, sentarse, estar de pie y acostarse. Durante la ejecución de estas actividades, se registraron datos del acelerómetro y giroscopio triaxial del dispositivo a una frecuencia de muestreo de 50 Hz.

II. DESCRIPCIÓN DEL PROBLEMA

A. Contexto del problema

El reconocimiento automático de actividades humanas a partir de datos de sensores móviles es un desafío fundamental para desarrollar aplicaciones inteligentes que mejoren la calidad de vida. Identificar con precisión qué actividad está realizando una persona, como caminar, sentarse o subir escaleras, permite crear sistemas personalizados para monitoreo de salud,

asistencia a personas mayores, control deportivo y entornos de interacción adaptativa.

Sin embargo, distinguir entre estas actividades usando únicamente datos crudos de acelerómetros y giroscopios resulta complejo debido a la variabilidad natural de los movimientos humanos y el ruido inherente a los sensores. Por ello, se requiere una solución capaz de extraer patrones significativos y generalizables que permitan clasificar correctamente cada actividad.

El uso de técnicas de *Machine Learning* (ML) ofrece un enfoque efectivo para resolver este problema, ya que permite construir modelos capaces de aprender automáticamente las características discriminativas de las actividades a partir de grandes volúmenes de datos. Estos modelos pueden adaptarse a diferentes usuarios y condiciones, ofreciendo precisión y robustez en entornos reales.

Desarrollar una solución basada en ML para el reconocimiento de actividades humanas facilita la creación de aplicaciones inteligentes con alto impacto social y comercial, tales como sistemas de rehabilitación remota, monitoreo continuo de salud, y dispositivos *wearables* que promuevan un estilo de vida saludable.

B. Composición de la base de datos

De acuerdo a la información asociada de [1], durante la recolección de datos, 30 voluntarios realizaron seis actividades mientras portaban un *smartphone* con sensores triaxiales (acelerómetro y giroscopio), muestreados a 50 Hz, es decir, 50 datos por segundo por eje. Las señales fueron segmentadas en ventanas de 2.56 segundos, equivalentes a 128 muestras por ventana. Para incrementar la cantidad de ejemplos y capturar mejor la variabilidad, se aplicó una superposición del 50% entre ventanas consecutivas, generando un desplazamiento de 1.28 segundos (64 muestras).

Este procedimiento generó una secuencia de ventanas etiquetadas de forma supervisada, asociadas tanto a la actividad realizada como al sujeto correspondiente. Como resultado, el conjunto de datos contiene 10,299 ventanas, cada una

representando un segmento temporal en el que se llevó a cabo una de las seis actividades predefinidas.

Las características corresponden a estadísticas calculadas sobre series temporales preprocesadas, tales como medias, desviaciones estándar, energías, magnitudes de señal, coeficientes de correlación, transformadas rápidas de Fourier (FFT), entre otras.

Las características se extraen de señales en los ejes X, Y y Z provenientes de dos fuentes: aceleración total y aceleración corporal (filtrada para eliminar la componente gravitacional), así como velocidad angular (giroscopio). Todas las señales fueron previamente normalizadas y transformadas para capturar tanto información temporal como frecuencial.

El dataset incluye además dos columnas meta: *subject*, que identifica al individuo que realizó la actividad (entero de 1 a 30), y *activity*, que representa la etiqueta de clase (una de seis categorías predefinidas).

La estructura se encuentra dividida en dos archivos principales:

- *train*, que contiene */X_train.txt*, */y_train.txt*, y */subject_train.txt*
- *test*, que contiene */X_test.txt*, */y_test.txt*, y */subject_test.txt*

C. Paradigma a utilizar

El paradigma a utilizar corresponde a un aprendizaje supervisado multiclase, donde existen diferentes muestras que representan las características de una actividad durante un intervalo de tiempo, tal que existe un total de 561 características procesadas para un intervalo de una actividad.

III. ESTADO DEL ARTE

Se realizó una exploración sobre diferentes estudios previos relacionados a HAR para entender diferentes perspectivas y soluciones aplicadas al problema. Entre ellas se destacan las siguientes:

- En [2] se desarrolló un modelo de *Support Vector Machines* (SVM) clasificador binario generalizado para casos multiclase mediante la aproximación *One-vs-All* (OVA), y se realizó la selección de los hiperparámetros a través de una validación cruzada de 10-fold y *kernels* Gaussianos. A través de la evaluación por *Precision* y *Recall* se obtiene un modelo con una precisión general del 96% para los datos de prueba, destacando que existen dificultades para la clasificación correcta entre *STANDING* y *SITTING* debido a la ubicación del sensor (cintura).
- En [3] se explora cómo las características físicas de los sujetos influyen en los modelos de inteligencia artificial utilizando una aproximación de "ensemble learning" con una multitud de algoritmos como *Random Forest*, *XGB*, *AdaBoost*, *Artificial Neural Network*, *Vanilla Recurrent Neural Network*, *Long Short Term Memory*, *KNN*, etc. Para el entrenamiento, proceden a agrupar los datos por sujetos de peso y altura similares, y posteriormente iteran clasificadores sobre el dataset construido utilizando validación cruzada de 10 Folds para asegurar resultados robustos, escogiendo aquel con mejores resultados de

Accuracy. Adicionalmente, realizan comparaciones con *Recall* (inexactitud del modelo respecto a valores positivos predichos como negativos), *Precision* (inexactitud del modelo respecto a valores negativos predichos como positivos) y *F1-Score* (media armónica entre *precision* y *recall*). Teniendo en cuenta que durante los entrenamientos .

Al final se encuentra que entrenar considerando los atributos físicos de los individuos ayuda a entrenar modelos más consistentes en sus predicciones y más resistentes a variaciones de *dataset*. Así mismo, los algoritmos que alcanzan el mayor *accuracy* son *Random Forest* y *XGB*; mientras que la familia de algoritmos que menos se beneficia son los algoritmos de *deep learning*, que en consideración de los autores tienden a sobreajustarse con facilidad y tienen menos capacidad para reconocer patrones estadísticos.

- En [4] se emplearon algoritmos como *Support Vector Machine* (SVM), *Random Forest*, *K-Nearest Neighbors* (k-NN) y redes neuronales profundas, que fueron entrenados y comparados en varios conjuntos de datos estándar, evaluando su desempeño a través de la validación cruzada entre dominios, que consistió en entrenar los modelos en un conjunto de datos y probarlos en otro diferente con el objetivo de medir su capacidad para generalizar y funcionar en nuevas condiciones sin necesidad de reentrenamiento.

Adicionalmente, el rendimiento se midió principalmente con la exactitud promedio en cinco ejecuciones por modelo, y también se utilizó la métrica *Maximum Mean Discrepancy* (MMD) para evaluar la diferencia estadística entre los conjuntos de datos. Al final, los resultados mostraron que el desempeño de los modelos disminuye notablemente al evaluarlos en datos de otro conjunto, incluso cuando la diferencia estadística es baja, debido a factores como la posición del dispositivo, el tipo de smartphone y características de los participantes. Además, las actividades de baja intensidad resultaron más estables entre dominios que las dinámicas.

- En [5] se busca mejorar la precisión del reconocimiento de actividades humanas usando *Exploratory Data Analysis* (EDA) junto con técnicas de reducción de dimensionalidad, intentando superar los resultados de los modelos tradicionales con altos costos computacionales y baja eficiencia. Para lograr este objetivo proponen utilizar un EDA que concentre su atención en la distribución real de los datos en conjunto con la técnica de reducción de la dimensionalidad *T-Distributed Stochastic Neighbor Embedding* (t-SNE) que permite la visualización de datos de alta dimensionalidad. .

Posteriormente, para los experimentos se utilizaron en los componentes de aceleración gravitacional y su posición en el cuerpo humano para realizar la clasificación a través de diferentes modelos como Regresión Logística, *LinearSVC*, *KernelSVM*, Árboles de decisión, *RandomForest* y *Gradient Boosting DT*. Así mismo, los parámetros de

cada modelo fueron optimizados a través de la técnica de validación cruzada *GridSearchCV* y la evaluación se realizó a través de las métricas *Accuracy*, *Recall*, *F1-score* y Matrices de confusión. Finalmente, se obtuvo que los mejores modelos fueron *LinearSVC* y *KernelSVM* cada uno con *accuracy* de 96,66% y 96,46% respectivamente, que corresponden a valores superiores cuando se comparan con otros estudios presentados en el artículo.

A partir de los experimentos realizados en otros artículos, podemos observar que existen factores clave en la obtención de buenos resultados usando modelos de inteligencia artificial para la clasificación de HAR, como son la posición de los sensores en el cuerpo humano, las características físicas del sujeto y la utilización de métodos de validación cruzada robustos.

IV. ENTRENAMIENTO Y EVALUACIÓN DE LOS MODELOS

A. Configuración Experimental

- 1) Metodología de Validación:
- 2) Configuración de Hiperparámetros:

- Modelo 1
- Modelo 2
- Modelo 3
- Modelo 4
- Modelo 5

TABLA I
TABLA DE CONFIGURACIÓN PARA LOS MODELOS

Configuraciones	Modelo 1			Modelo 2			Modelo 3			Modelo 4			Modelo 5		
	HP1	HP2	HP3	HP1	HP2	HP3	HP1	HP2	HP3	HP1	HP2	HP3	HP1	HP2	HP3
Config 1	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0
Config 2	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0
Config 3	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0
Config 4	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0
Config 5	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0

- 3) Métricas de Desempeño:

B. Resultados del entrenamiento de modelos

Luego de realizar el entrenamiento y las pruebas de cada uno de los modelos, se obtuvieron los siguientes resultados:

V. REDUCCIÓN DE DIMENSIÓN

A. Selección de características

B. Extracción de características

VI. CONCLUSIONES

A partir de lo realizado se puede concluir que ...

TABLA II
RESULTADOS DE ENTRENAMIENTO POR MODELO

Modelo	Entrenamiento	Validación	Pruebas
Modelo 1	0	0	0
Modelo 2	0	0	0
Modelo 3	0	0	0
Modelo 4	0	0	0
Modelo 5	0	0	0

REFERENCIAS

- [1] J. Reyes-Ortiz, D. Anguita, A. Ghio, L. Oneto, and X. Parra. "Human activity recognition using smartphones", UCI Machine Learning Repository, 2013. [En línea]. Disponible: <https://doi.org/10.24432/C54S4K>.
- [2] D. Anguita, A. Ghio, L. Oneto, X. Parra, y J. L. Reyes-Ortiz, "A Public Domain Dataset for Human Activity Recognition using Smartphones", The European Symposium on Artificial Neural Networks, 2013.
- [3] N. A. Choudhury, S. Moulik, y D. S. Roy, "Physique-Based Human Activity Recognition Using Ensemble Learning and Smartphone Sensors", IEEE Sensors Journal, vol. 21, no. 15, pp. 16852-16860, 2021. [En línea]. Disponible: <https://doi.org/10.1109/JSEN.2021.3077563>
- [4] O. Napoli, D. Duarte, P. Alves, et al., "A benchmark for domain adaptation and generalization in smartphone-based human activity recognition", Sci Data, vol. 11, p. 1192, 2024. [En línea] Disponible: <https://doi.org/10.1038/s41597-024-03951-4>
- [5] W. Kong, L. He, y H. Wang, "Exploratory Data Analysis of Human Activity Recognition Based on Smart Phone", IEEE Access, vol. 9, pp. 73355-73364, 2021. [En línea]. Disponible: <https://doi.org/10.1109/ACCESS.2021.3079434>