

Entrega 1

Contexto de aplicación

El presente proyecto utilizará datos obtenidos de un experimento que incluyó a 30 voluntarios, con edades entre 19 y 48 años. Cada voluntario portaba un smartphone Samsung Galaxy S II, con el cual se registraron datos de aceleración lineal tridimensional y velocidad angular tridimensional. Estos datos se capturaron a una tasa de muestreo constante de 50 Hz, empleando el acelerómetro y el giroscopio incorporados en el dispositivo [1].

Las actividades realizadas por los participantes se dividieron en seis categorías distintas: caminar (WALKING), subir escaleras (WALKING_UPSTAIRS), bajar escaleras (WALKING_DOWNSTAIRS), estar sentado (SITTING), estar de pie (STANDING) y estar acostado (LAYING).

El proyecto se propone aplicar técnicas de Deep Learning, específicamente de modelos secuenciales, para lograr una identificación precisa de las actividades humanas con base en el análisis de las señales de aceleración y giroscopio proporcionadas por el smartphone.

Objetivo de machine learning

El objetivo es identificar la actividad realizada por una persona (Caminar, Subir escaleras, Bajar escaleras, Sentarse, Ponerse de pie, Acostarse) basándose en las series temporales de las señales inerciales, utilizando redes neuronales convolucionales (CNN), redes neuronales recurrentes (RNN), Long Short-Term Memory (LSTM), y Gated Recurrent Unit (GRU) con el fin de determinar cuál presenta un mejor rendimiento en la detección de las actividades.

Dataset:

Las señales del sensor (acelerómetro y giroscopio) fueron preprocesadas aplicando filtros de ruido y luego muestreadas en ventanas deslizantes de ancho fijo de 2.56 segundos con un solapamiento del 50% (128 lecturas/ventana). La señal de aceleración del sensor, que tiene componentes de gravedad y movimiento del cuerpo, fue separada usando un filtro pasa bajos Butterworth en aceleración del cuerpo y gravedad. Se asume que la fuerza gravitatoria solo tiene componentes de baja frecuencia, por lo tanto, se utilizó un filtro con una frecuencia de corte de 0.3 Hz.

Tamaño:

Número de registros: Se cuenta con un total de 7352 registros de entrenamiento y 2947 de prueba.

Longitud de registro: Cada registro cuenta con 128 lecturas.

Tipo de Dato: Series temporales.

Número de Características: 9, aceleración triaxial del cuerpo, aceleración angular triaxial, aceleración total triaxial.

Tamaño en disco: Se necesita un espacio de aproximadamente 129.6MB para los datos de entrenamiento y 51.84 MB para datos de prueba.

Distribución de Clases: El dataset incluye seis clases que representan diferentes actividades: WALKING, WALKING_UPSTAIRS, WALKING_DOWNSTAIRS, SITTING, STANDING, LAYING. La distribución de clases es equitativa, con cada clase representada aproximadamente en la misma proporción.

Métricas de desempeño de Machine Learning.

Accuracy: Proporción de predicciones correctas respecto al total de predicciones.

$$\text{Precisión} = \frac{\text{Número de predicciones correctas}}{\text{Número total de predicciones}}$$

Recall: Proporción de predicciones positivas correctas respecto a todas las verdaderas positivas.

$$\text{Recall} = \frac{\text{True Positives}}{\text{True Positives} + \text{False Negatives}}$$

Precisión: Número de elementos identificados correctamente como positivo de un total de elementos identificados como positivos.

$$\text{Precisión} = \frac{\text{True Positives}}{\text{True Positives} + \text{False Positives}}$$

F1-Score: Media armónica de precisión y recall, proporciona un equilibrio entre ambas métricas.

$$F1 - \text{Score} = 2 \times \frac{\text{Precisión} \times \text{Recall}}{\text{Precisión} + \text{Recall}}$$

Matriz de Confusión: Proporciona información sobre los aciertos y errores de clasificación para cada clase.

Métricas de Desempeño de Negocio:

Tiempo de Inferencia:

El tiempo que tarda el modelo en hacer una predicción una vez que recibe nuevos datos de entrada.

Rentabilidad:

Estimación del impacto financiero del modelo en el contexto de la bioingeniería.

Resultados previos

Arora et al [2] aplicaron redes neuronales de tipo LSTM al conjunto de datos. El modelo propuesto consta de dos capas LSTM y una capa de dropout con un porcentaje del 50% para

mitigar el sobreajuste. Se logró una precisión del 92.14%, superando a modelos de aprendizaje clásicos. Aunque la clasificación fue robusta para la mayoría de las actividades, se observó cierta dificultad en distinguir entre sentarse y estar de pie. Se planteó la posibilidad de explorar modelos más avanzados, optimizar hiperparámetros y seleccionar críticamente el tamaño del lote (batch size), las épocas (epochs) y el porcentaje de dropout para mejorar el rendimiento de la clasificación.

Ali et al [3] condujeron un estudio centrado en el reconocimiento de actividades humanas (HAR), se empleó una metodología basada en redes neuronales convolucionales (CNN) para analizar los datos de interés. La preparación de datos incluyó la transformación de series temporales tridimensionales en imágenes para ser procesadas por la CNN. La metodología incluyó entrenamiento, validación y evaluación del modelo, con parámetros clave como el número de capas, tasa de aprendizaje y tamaño del lote controlados. Se logró una alta precisión de clasificación, alcanzando 99.8% en las pruebas.

Ortiz et al [4] implementaron la base de datos objeto de estudio, la cual es de su autoría, con el fin de construir y evaluar una arquitectura de Reconocimiento de Actividad Humana Sensible a Transiciones (TAHAR) para lo cual combinan la salida probabilística de predicciones de actividades consecutivas de una Máquina de Vectores de Soporte (SVM) con un enfoque de filtrado heurístico.

En general, el dataset a implementar ha sido ampliamente usado en el campo de Deep Learning, sin embargo sus enfoques están orientados al uso de las características previamente extraídas de las series temporales para resolver problemas de clasificación HAR entre otras aplicaciones, por lo que será interesante explorar la clasificación haciendo uso directamente de las series temporales.

Referencias

- [1] D. Anguita, A. Ghio, L. Oneto, J.-L. Parra, Xavier.and Reyes-Ortiz, A public domain dataset for human activity recognition using smartphones, in: European Symposium On Artificial Neural Networks, Computational Intelligence and Machine Learning ESANN, 2013.
- [2] G. Arora, A. Ahlawat, and M. Payal, “Long Short Term Memory network for Recognition of Daily Human Activity,” Mar. 2021. doi: 10.4108/eai.27-2-2020.2303125.
- [3] G. Q. Ali and H. Al-Libawy, ‘Time-Series Deep-Learning Classifier for Human Activity Recognition Based On Smartphone Built-in Sensors’, *Journal of Physics: Conference Series*, vol. 1973, no. 1, p. 012127, Aug. 2021.
- [4] Reyes Ortiz, Jorge Luis ; Oneto, Luca ; Samà Monsonís, Albert ; Ghio, Alessandro ; Llanas Parra, Xavier ; Anguita, Davide. "Transition-aware human activity recognition using smartphones". Universitat Politècnica de Catalunya. Departament d'Enginyeria Electrònica.Marz. 2015.