

## Entrega 1

### Preguntas de Interés

- ¿Qué características de los movimientos articulares son más relevantes para clasificar actividades humanas?
- ¿Qué técnicas de preprocesamiento son más efectivas para mejorar la calidad del seguimiento de las articulaciones?
- ¿Cuál es el mejor modelo de clasificación (de los vistos en clase) para predecir actividades humanas en tiempo real?
- ¿Cuáles son las mejores métricas de evaluación para asegurar que el sistema sea útil y preciso en tiempo real?
- ¿Qué nivel de precisión (u otras métricas) es necesario para que el sistema sea útil en aplicaciones del mundo real?

### Tipo de problema

Este es un problema de clasificación multiclase supervisada.

### Metodología

Para nuestro proyecto vamos a seguir la metodología ASUM-DM, que es una versión mejorada de CRISP-DM, ya que consideramos que es adecuada para proyectos como el nuestro, donde tenemos que manejar muchos datos y realizar varias iteraciones para mejorar los datos y los modelos. Esta metodología nos ayudará a organizar mejor el trabajo, permitiendo automatizar algunas tareas y reutilizar lo que ya hayamos hecho en cada fase (como la comprensión de los datos, preparación, modelado y evaluación). Esto nos permitirá iterar más rápido y ajustar el sistema de seguimiento de movimientos en tiempo real según vayamos obteniendo resultados, haciendo que todo el proceso sea más ágil y eficiente.

### Herramientas

Las herramientas propuestas para usarse durante el proyecto son las siguientes:

- **Jupyter Notebook** para la creación y documentación del proyecto, permitiendo combinar código, visualizaciones y texto.
- **OpenCV** y **MediaPipe** proporcionan capacidades avanzadas de procesamiento de imágenes y detección de poses.
- **NumPy** y **Pandas** manejan y manipulan los datos de manera eficiente, preparando los conjuntos de datos para el análisis y modelado.
- **Matplotlib** y **Seaborn** se encargan de la visualización de los datos y los resultados para la interpretación de la información.

- **Scikit-learn** permite implementar y evaluar modelos de aprendizaje supervisado y no supervisado automáticos.

## Métricas

Usaremos las siguientes métricas para evaluar nuestros modelos de clasificación:

- **Accuracy:** nos muestra el porcentaje de predicciones correctas del total, es útil para evaluar el desempeño general del modelo.
- **Precisión:** indica el porcentaje de predicciones positivas correctas sobre todas las predicciones positivas, es útil para minimizar los falsos positivos.
- **Recall:** Mide el porcentaje de casos positivos que fueron correctamente identificados por el modelo, útil para minimizar los falsos negativos.
- **F1 Score:** es útil para hacer un balance entre precisión y recall.
- **Matriz de Confusión:** nos permite visualizar el rendimiento del modelo clasificando correctamente o incorrectamente cada clase.
- **ROC:** Nos muestra el rendimiento del modelo en diferentes umbrales de clasificación, es útil para evaluar la relación entre la tasa de verdaderos positivos y la tasa de falsos positivos.

## Datos

- El conjunto inicial de datos recolectados fueron un total de 6 videos grabados con la ayuda de una compañera. En los videos se mezclan actividades como acercarse, alejarse, girar y sentarse, de tal forma que puedan ser usados y procesados durante el análisis exploratorio de datos para ser de utilidad en el entrenamiento del futuro modelo.
- Tomando en cuenta que la clasificación de movimiento va sujeta a otras variables como lo son los aspectos físicos de la persona (altura, ancho, entre otros), la posición de la cámara, la iluminación y objetos que puedan interferir en la captura. Se hace necesario entonces continuar la grabación de videos en los que se visualicen las acciones y tomar en consideración el uso de datasets como lo son NTU RGB+D, Kinetics Dataset, UCF101, HMDB que agrupan muestras de videos entre las que se encuentran algunas de las etiquetas del proyecto, que bajo la licencia para investigación que ofrecen serían de gran utilidad.

## EDA:

Tras la obtención de los videos se hizo uso de las herramientas Media Pipe y OpenCV para llevar a cabo el seguimiento de la postura de diferentes puntos a lo largo del cuerpo. La cantidad elegida fueron 33 ya que según múltiples fuentes, esta cantidad permite llevar a cabo un rastreo adecuado y preciso de todo el cuerpo.

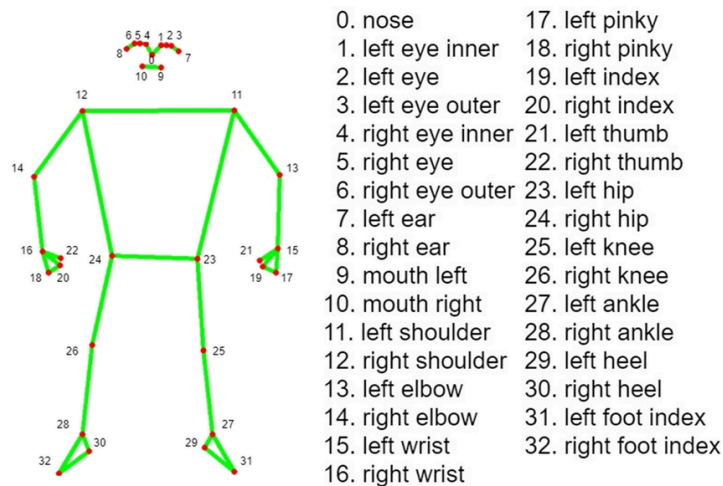
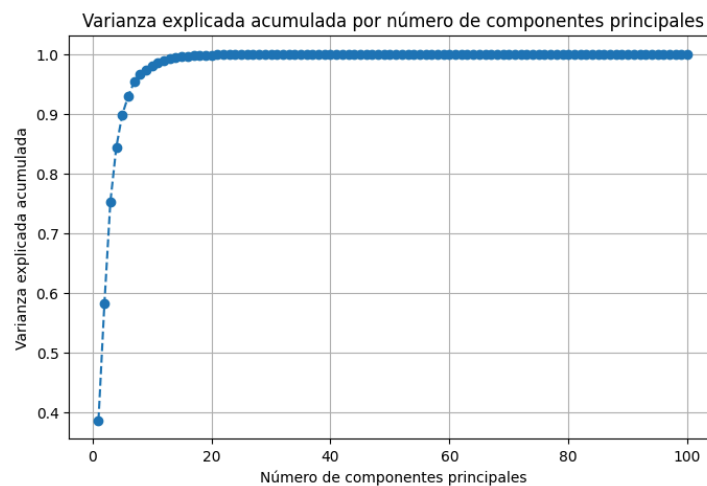


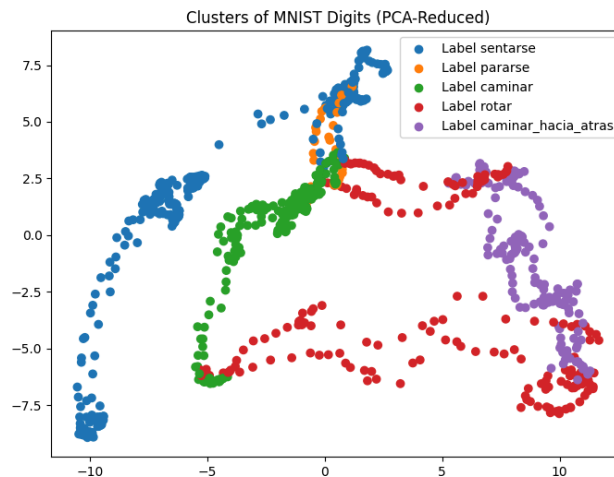
Figure - available from: Journal of Ambient Intelligence and Humanized Computing

De esta forma logramos transformar cada frame del video en un registro con 33 landmarks, cada una con su respectivo X, Y y Z, además del label correspondiente a la actividad.

Para entender la distribución de los datos y ver si es posible diferenciar grupos claros desde un inicio, decidimos realizar una gráfica de varianza acumulada por componentes, para ver a cuantos componentes podemos reducir los datos sin perder información significativa:



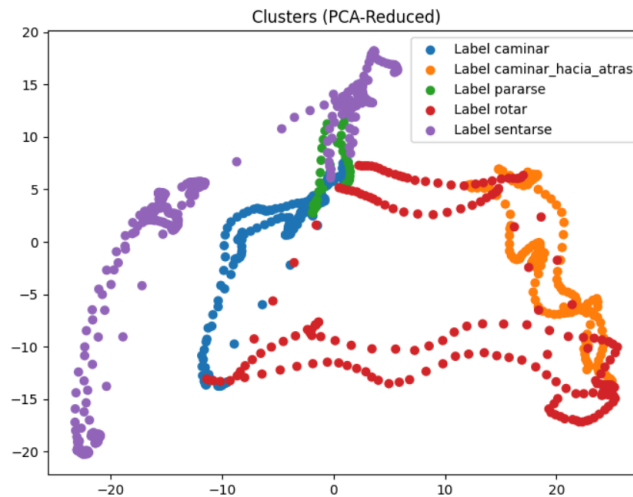
Con esta información se pudo identificar que con 8 componentes es suficiente para explicar el 95% de los datos, así que realizamos la reducción de componentes con ayuda de PCA y graficamos la distribución de los datos:



En esta gráfica se pueden observar los cinco grupos de movimientos (cada uno con un color) y en su mayoría son claramente diferenciables, por lo que los modelos que se implementen posteriormente deberían poder hacer un buen trabajo; sin embargo, se identificó un problema con los datos con la etiqueta “pararse”, pues se pierden casi por completo entre los movimientos de sentarse, y caminar hacia adelante. Pensamos que esto se puede deber a que cada uno de los puntos representados es un único fotograma, por lo que no hay realmente una secuencia o contexto entre los fotogramas, y dos acciones como sentarse y pararse tendrán las mismas imágenes en diferente orden.

Inicialmente optamos por agrupar las fotogramas con una misma etiqueta en grupos de 5 de la siguiente forma: 1, 2, 3, 4, 5 en la primera fila, 6, 7, 8, 9, 10 en la segunda y así hasta el final. Sin embargo al hacer uso de este enfoque la cantidad de datos se ve reducida en 5 veces su tamaño, lo que podría afectar negativamente el entrenamiento y análisis puesto que si del frame 1 al 10 la acción es la misma la unión 3, 4, 5, 6, 7 también debería estar presente en los datos de entrenamiento.

Para solucionarlo decidimos agrupar nuevamente los fotogramas que tienen la misma etiqueta en grupos de 5, pero de la siguiente forma: los fotogramas 1,2,3,4,5 se juntan en la primera fila, los fotogramas 2,3,4,5,6 se juntan en la segunda fila y así hasta el final, de tal forma que cada frame conoce los próximos 5 y no se ve afectada en gran medida la cantidad de datos de entrenamiento, lo que teorizamos tendrá un efecto positivo en el futuro modelo. El resultado, después de identificar que necesitamos usar 20 componentes, es el siguiente:



Se observa una pequeña mejoría, pero sigue sin ser suficiente para que un modelo de clasificación pueda diferenciar bien las acciones, por lo que consideramos que debemos tomar más videos para mejorar el conjunto de datos.

### Problema ético:

Al desarrollar nuestro proyecto, es fundamental considerar varios aspectos éticos. En primer lugar, la privacidad de los videos de entrenamiento es muy importante, ya que estos capturan a personas realizando actividades específicas, es esencial obtener su consentimiento informado y garantizar que los datos sean almacenados y utilizados de manera segura. Además, debemos ser conscientes de los riesgos de discriminación en el entrenamiento del modelo; si los videos no incluyen una representación diversa en términos de género, complexión física o capacidades motrices, el modelo podría generar sesgos y realizar predicciones menos precisas para ciertos grupos, lo que lo haría poco inclusivo.

También es crucial considerar el uso inadecuado del modelo para vigilancia sin consentimiento. Aunque el propósito de nuestra herramienta es el seguimiento postural, podría ser mal utilizada para monitorear las acciones de alguien sin su aprobación, lo que plantea serias preocupaciones sobre la privacidad por lo que debemos establecer límites claros en el uso del sistema.

Finalmente, si el modelo se aplica en procedimientos críticos donde los errores no son aceptables, debemos garantizar que haya sido rigurosamente probado y que se implementen medidas de seguridad para minimizar riesgos. Es fundamental ser transparentes sobre sus limitaciones y asegurarnos de que los usuarios comprendan el contexto en el que se usa.

### Siguientes pasos

Para avanzar en nuestro proyecto tenemos varias etapas que debemos seguir:

1. **Reducción de dimensionalidad:** Antes del entrenamiento y mientras el conjunto de datos siga en crecimiento aplicaremos técnicas de reducción de dimensionalidad

como PCA para simplificar los datos y eliminar ruido. Esto puede mejorar el rendimiento del modelo al enfocarse en las características más significativas de los datos.

2. **Entrenamiento del modelo:** entrenaremos nuestro modelo de clasificación utilizando los datos de entrenamiento recopilados, y seleccionaremos los algoritmos adecuados, como SVM, Random Forest o XGBoost, para comparar y validar los resultados e intentar llegar al mejor modelo.
3. **Ajuste de Hiper-Parámetros:** utilizaremos técnicas como la validación cruzada y grid search para encontrar la mejor combinación de hiper-parámetros que maximicen las métricas y optimicen el rendimiento
4. **Implementación del sistema en tiempo real:** esto implica integrar el modelo entrenado con una interfaz gráfica que permita a los usuarios observar la clasificación de actividades y el seguimiento de movimientos articulares en tiempo real.
5. **Continuar recolectando más datos:** planeamos realizar grabaciones adicionales que ayuden a enriquecer nuestro conjunto de datos y abordar posibles sesgos que se hayan identificado durante el entrenamiento inicial.