Project 1:

Data Analysis for medical applications.

Instituto Tecnológico y de Estudios Superiores de Monterrey

Campus Guadalajara

Marissa Edith Luna Landa A01645653

Ximena Cantón Ayllón A01645105

Modeling Learning with Artificial Intelligence

Omar Mendoza Montoya

Group 302

Date: Viernes, 06 de junio de 2025.

1. Description of the implemented application and the collected data

What does the application do?

The developed application aims to classify, in real time, one of six different physical activities performed by a user, using data from a mobile device's accelerometer. The activities recognized by the model are:

- Nothing (rest)
- Jumping
- Lateral lunges
- Hip circles
- Walking straight
- Squats

The system was previously trained with data manually collected by team members, performing each activity during defined time intervals. Later, the trained model was integrated into an application that receives new input data and predicts in real time which activity is taking place.

What is it expected to do?

The application is expected to receive accelerometer signals from the device (in this case, from a smartphone using the Phyphox app). Once the signals are received, the model extracts the most relevant features from the observations and classifies the activity in real time using a trained Random Forest model with optimized parameters; aiming for the highest possible prediction accuracy.

In essence, the application enables an automated system to recognize the type of physical movement a user is performing, without manual intervention

What are some examples of similar applications?

Human Activity Recognition (HAR) is a well-established research area that already has practical applications in fields such as fitness, healthcare, and personalized assistance.

For example, platforms like **Google Fit**, **Apple Health**, and wearable devices such as **Fitbit**, **Garmin**, or **Polar** use inertial sensors (accelerometers and gyroscopes) to identify whether a person is walking, running, cycling, or resting. These tools not only recognize activities but also enable continuous monitoring of physical condition, calorie estimation, active time tracking, and user feedback (Machorro-Cano et al., 2023).

Moreover, in the **medical field**, HAR systems have been used to support physical rehabilitation processes. There are applications that monitor whether a patient correctly performs prescribed exercises, even remotely. This has proven especially useful for patients with limited mobility, such as those recovering from **stroke incidents** (Chen et al., 2016).

The integration of wearable sensors allows for objective and automated tracking of patient progress.

Additionally, **AI-integrated HAR frameworks** have been developed to combine physiological, environmental, and inertial data, enhancing system accuracy and adaptability across different contexts (Demrozi et al., 2020). This opens opportunities for use in **smart homes**, **elderly monitoring**, and **autonomous assistance systems** (Kaur & Kaur, 2024; Liu et al., 2021).

Our project follows a similar line: we designed an application capable of recognizing specific physical activities from accelerometer signals, training models that can run online and adapt to different users. Such a system has strong potential in **sports environments**, **rehabilitation**, and **remote monitoring**, as it enables real-time detection of whether certain movements are being performed correctly.

What type of information will it monitor?

The application is designed to monitor **real-time body movement** using the accelerometer of a mobile device. Specifically, it records acceleration values along the three spatial axes: **X** (left–right), **Y** (forward–backward), and **Z** (up–down). From these signals, various **statistical features** are extracted to summarize movement behavior over short time intervals.

To structure the data collection process, **10 trials** were defined for each activity, and within each trial, **10 analysis windows** of **0.5 seconds** each were generated, with a **sampling rate of 30 Hz**. This resulted in **100 observations per activity**, and since **6 different activities** were recorded in each trial, a total of **600 observations per participant** were obtained.

For each window, multiple features were calculated—such as **mean, standard deviation, kurtosis, skewness, maximum, minimum, peak-to-peak range, derivatives**, among others. These features form an **input vector** used to feed the classification models.

The ultimate goal of the application is to **automatically recognize which physical activity** the person is performing based on the captured acceleration patterns, enabling potential applications in both **sports** and **clinical** contexts.

2. Descripción de las características extraídas de los datos.

From the original data, additional metrics were derived and added as features in the dataset. This process is necessary because **raw signals are more difficult to interpret** compared to statistical features, which help explain how and why a model makes certain decisions (Fulcher & Jones, 2014).

The goal of these features is to capture different aspects of movement—such as **intensity, variability, symmetry,** or **abruptness**—so that the model can effectively distinguish between similar activities.

The features extracted for each axis (**X, Y, Z**) were:

- **Mean:** average acceleration within the window.
- **Standard deviation (std):** variability of the signal.
- **Kurtosis:** presence of extreme values.
- **Skewness:** asymmetry or bias of the signal.
- **Minimum and maximum:** extreme acceleration values.
- **Median:** central value of the distribution.
- **25th and 75th percentiles:** values describing the interquartile dispersion.
- **Peak-to-peak (ptp):** difference between maximum and minimum values.
- **Mean and standard deviation of the derivative:** estimation of acceleration change (smoothness or abruptness of movement).

Additionally, **combined features** were included, such as the **total magnitude of the acceleration vector**, calculated as:

$$\sqrt{X^2 + Y^2 + z^2}$$

In total, 37 features were extracted per window, allowing the construction of representative input vectors for the machine learning model.

3. Results obtained from the evaluation of the tested classifiers, including hyperparameter tuning and feature selection.

A total of **ten different classifiers** were tested. Below is a summary sheet for each model showing the results obtained.

Accuracy, precision, and recall:

- Linear SVM

  Accuracy: 0.89

Class	Precision	Recall
1	0.93	0.93
2	0.85	0.89
3	0.88	0.82
4	0.85	0.87
5	0.96	0.98
6	0.88	0.86

- RBF SVM

  Accuracy:

Class	Precision	Recall
1	1.00	0.94
2	0.86	0.83
3	0.95	0.76
4	0.72	0.90
5	0.91	0.98
6	0.83	0.78

- LDA

  Accuracy: 0.83

Class	Precision	Recall
1	1.00	0.93
2	0.84	0.71
3	0.97	0.72
4	0.70	0.85
5	0.86	0.98
6	0.71	0.78

- K-NN

  Accuracy: 0.91

Class	Precision	Recall
1	1.00	0.96
2	0.96	0.93
3	0.98	0.84
4	0.79	0.94
5	0.94	0.97
6	0.88	0.86

- Two layer MLP

  Accuracy: 0.91

Class	Precision	Recall
1	0.97	0.95
2	0.94	0.94
3	0.94	0.88
4	0.85	0.90
5	0.94	0.96
6	0.88	0.87

- Random Forest

  Accuracy : 0.93

Class	Precision	Recall
1	0.99	0.96
2	0.96	0.95
3	0.97	0.90
4	0.85	.91
5	0.94	0.97
6	0.90	0.89

- Naive Bayes

  Accuracy (Exactitud): 0.89

Class	Precision	Recall
1	0.99	0.94
2	0.93	0.91
3	0.91	0.79
4	0.81	0.89
5	0.95	0.97

6	0.80	0.86

- Extra trees

  Accuracy: 0.82

Class	Precision	Recall
1	0.94	0.91
2	0.81	0.79
3	0.76	0.80
4	0.74	0.79
5	0.95	0.92
6	0.75	0.72

- Gradient Boosting

  Accuracy: 0.92

Class	Precision	Recall
1	0.99	0.94
2	0.93	0.95
3	0.96	0.90
4	0.84	0.90
5	0.94	0.96
6	0.92	0.90

- Ridge Classifier

  Accuracy: 0.77

Class	Precision	Recall
1	1.00	0.90
2	0.79	0.70
3	0.90	0.77
4	.80	0.59

5	0.65	0.99
6	0.69	0.72

Hyperparameter tuning and feature selection:

- **Linear SVM**

Optimized hyperparameter: regularization coefficient (**C**).

The optimal value of **C** was found using **GridSearchCV()**. In the following graph, it can be observed that as **C** increases, the model's accuracy also improves—up to a point where it stabilizes and the accuracy remains within very similar ranges.

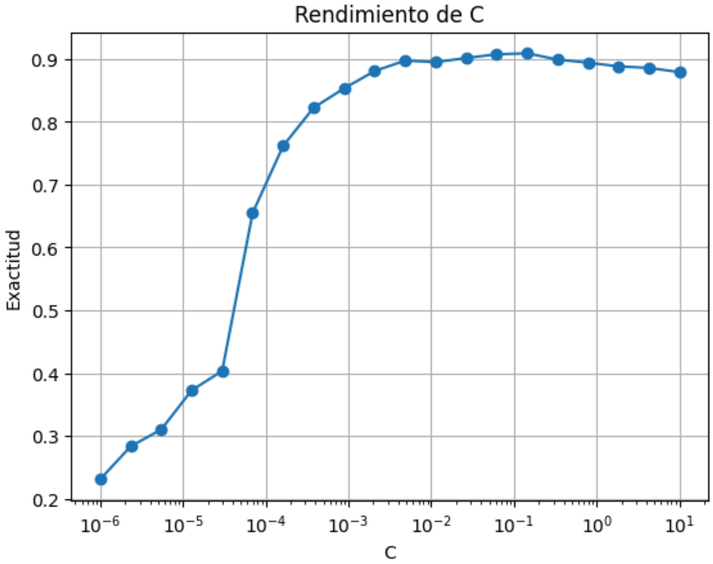

Fig 1. Performance of C for the Linear SVM model

Features used: Out of the 37 extracted features, cross-validation was performed to evaluate different feature subsets obtained through Recursive Feature Elimination (RFE).

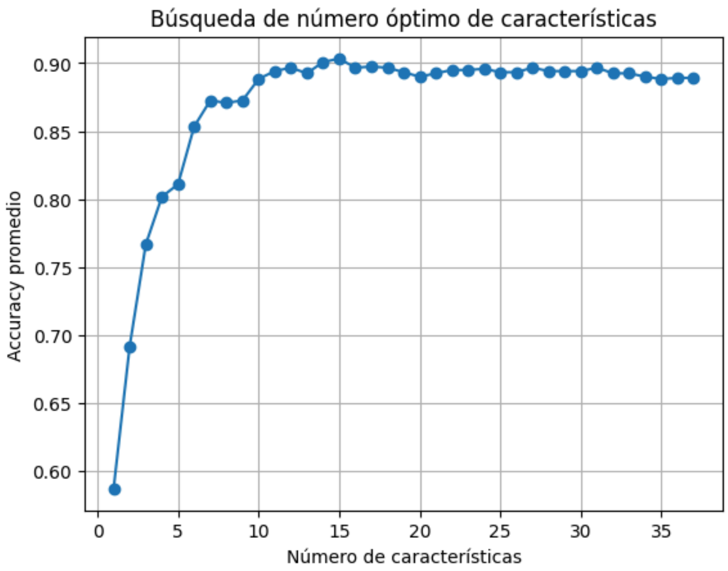

Fig 2. Optimal number of features for the Linear SVM model.

- **Random Forest**

Hyperparameters used: number of trees (**n_estimators**).

The accuracy performance of the **n_estimators** hyperparameter was plotted. In the following figure, it can be observed that within the tested range of trees, accuracy remains consistently high (between **0.90 and 0.94**).

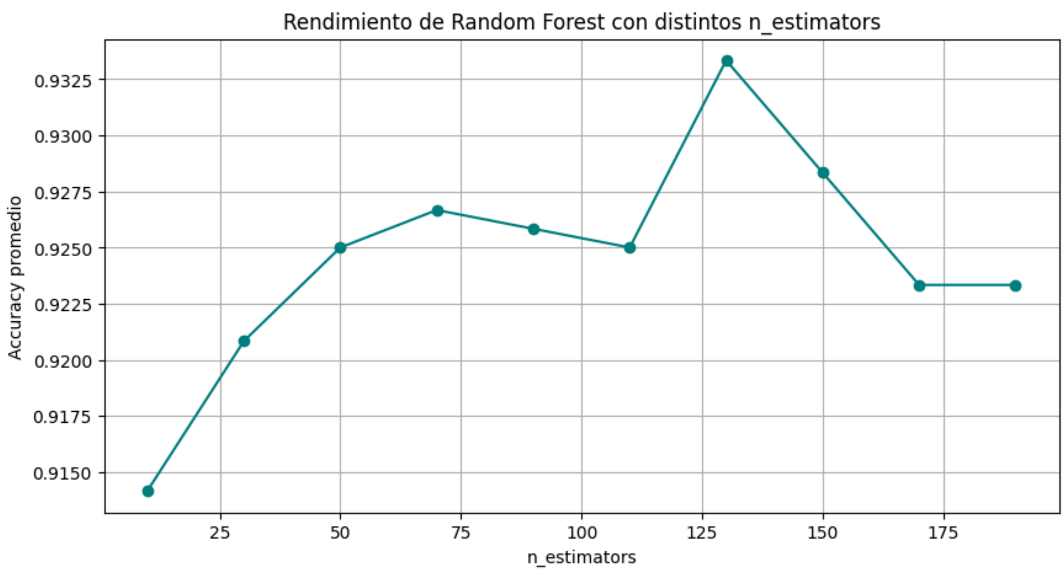

Fig 3. Performance of n_estimators for the Random Forest model.

Features used: RFE was again applied for feature selection, revealing that from 18 features onward, the model's accuracy stabilizes.

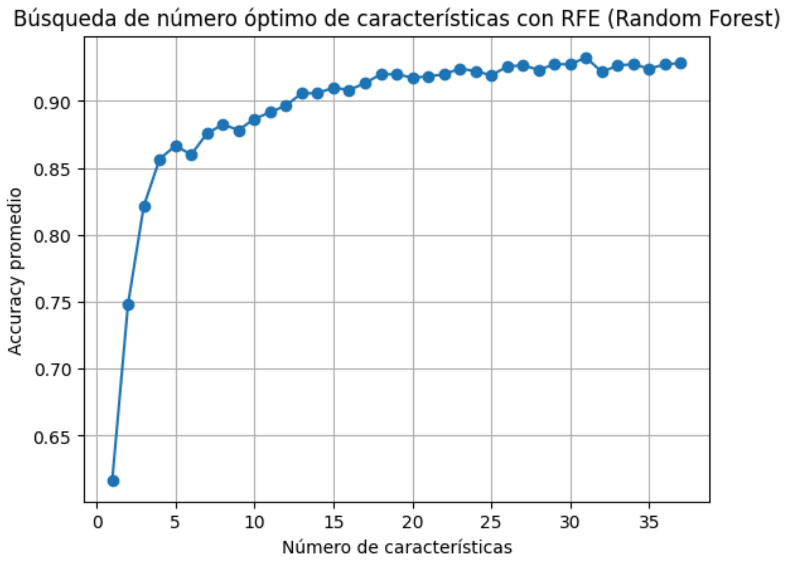

Fig 4. Optimal number of features for Random Forest.

4. Results of the online application and responses to the following questions:

Does it work the same for all team members data?

Initially, separate models were trained for each team member, and it was observed that the real-time classification performance varied noticeably depending on who performed the activity. This suggests that the model was learning **individual-specific movement patterns**, such as intensity or rhythm.

Subsequently, the data were unified into a **single training dataset** containing records from all participants. Although this led to a slight decrease in average accuracy, it achieved **greater generalization**: the model performed more consistently across all team members, regardless of who carried out the activity.

Is the online application performance consistent with the cross-validation results?
Yes. The performance observed in the online application was consistent with the results obtained during cross-validation. Activities were recognized accurately in real time; particularly those with more defined patterns, such as **jumping** or **walking straight**. On the other hand, **hip circles** proved to be one of the most challenging to predict in practice, which aligned with the cross-validation results, where this activity showed the **lowest precision and recall**.

Minor performance differences may be attributed to variations in how the activities were executed or to signal noise, but overall, the model responded **reliably and stably**.

5. Personal Conclusions

Marissa:

This project helped me see machine learning models from a completely new perspective. It's very easy to just train a model and stop there, but now I understand that it's much more than that. It's a whole system that needs to be understood, evaluated, and fine-tuned, and every part of that process has a direct impact on how well it performs.

One of my biggest takeaways is the importance of **cross-validation**. This technique is not only useful to estimate overall performance but also to evaluate by class. A model can have high overall accuracy while still failing on specific classes, which becomes a serious issue if those classes are critical in a real-world application.

Another key insight for me was realizing that **hyperparameter tuning** and **feature selection** are not mechanical tasks, they are decisions that should be guided by a real understanding of the model. Knowing how parameters like **C** in SVM or **n_estimators** in Random Forest affect performance allows for more thoughtful optimization rather than just random trial and error. Learning about feature selection was also eye-opening; it showed me how to reduce dimensionality, focus on what truly matters, save computational time, and at the same time, improve both the interpretability and efficiency of the model.

Overall, this project gave me a more complete view of a model's **life cycle**: from data collection and preprocessing to its implementation in a functional online application. I'm walking away with valuable tools, but more importantly, with the understanding that it's not just about using models, but about knowing **how and why they work** in order to truly improve them.

Ximena:

This project allowed me to understand machine learning model development from a much more comprehensive perspective. Beyond simply training a model, I realized that the entire system: from **data acquisition and preprocessing to validation and online deployment,** has a direct impact on final performance.

Something that truly stood out to me was recognizing that a model can show acceptable overall metrics and still fail on the most important classes. This helped me appreciate the role of **cross-validation**, not just as a way to measure global accuracy, but as a key tool for identifying specific errors and making more informed decisions.

I also experienced firsthand the importance of understanding **hyperparameters** and **feature selection**. It's not just about finding the best combination; it's about knowing how and why those elements influence model behavior. That kind of understanding gave me much greater clarity when tuning and optimizing performance.

Finally, deploying the model in a **working online application** helped me connect all the stages of the process and appreciate the value of each one. I'm leaving with a broader and more practical view of what it means to design an intelligent system: it's not enough for the model to simply "work"; it's just as important to understand its **logic, limitations, and how to make it better**.

6. References

Chen, K., Li, Y., & Zhou, W. (2016). *Evaluation of a smartphone human activity recognition application with able-bodied and stroke participants*. Journal of NeuroEngineering and Rehabilitation, 13(1), 1–10. https://jneuroengrehab.biomedcentral.com/articles/10.1186/s12984-016-0114-0

Demrozi, F., Pravadelli, G., Bihorac, A., & Rashidi, P. (2020). *Human Activity Recognition using Inertial, Physiological and Environmental Sensors: A Comprehensive Survey*. arXiv preprint arXiv:2004.08821. https://ieeexplore.ieee.org/document/9257355

Fulcher, B. D., & Jones, N. S. (2014). Highly comparative feature-based time-series classification. *IEEE Transactions on Knowledge and Data Engineering*, 26(12), 3026–3037. https://doi.org/10.1109/TKDE.2014.2316504

Kaur, P., & Kaur, P. (2024). *Human activity recognition: A comprehensive review*. Expert Systems, 41(1), e13680. https://onlinelibrary.wiley.com/doi/10.1111/exsy.13680

Liu, R., Ramli, A. A., Zhang, H., Henricson, E., & Liu, X. (2021). *An overview of human activity recognition using wearable sensors: Healthcare and artificial intelligence*. arXiv preprint arXiv:2103.15990. https://link.springer.com/chapter/10.1007/978-3-030-96068-1_1

Machorro-Cano, I., Olmedo-Aguirre, J. O., Alor-Hernández, G., Rodríguez-Mazahua, L., Sánchez-Morales, L. N., & Pérez-Castro, N. (2023). *Cloud-based platforms for health monitoring: A review*. Informatics, 10(1), 1–20. https://www.mdpi.com/2227-9709/10/1/3