**REQUIREMENTS FOR DATASET**

**Description**: The dataset must contain information about all training activities registered by a smartwatch or smart-application. The data includes information about the type of activity, the elapsed time, the duration, the calories, and other related information.

**Data format:** The data must be stored in a CSV file.

**Data size:** The size of the dataset must be as big as possible.

**Required data fields:**

The dataset must have the fields in bold of the following table. The others are also accepted, but not mandatory.

The header names must be in CamelCase format, exactly the same as the following table:

|  |  |
| --- | --- |
| **ActivityType** | The type of activity must be one of the followings: Mountain biking, Road biking, Spinning, Weight training, Running, Athletic Walking, Alpinism, Swimming, Hike or Walk.  Must be written either in Englis or in Spanish. |
| **Date** | The date on which the activity took place.  yyyy-mm-dd hh:mm:ss |
| **Title** | The title of the activity. |
| **Distance** | The distance covered during the activity. Units must be Km |
| **Calories** | The number of calories burned during the activity. |
| **Time** | The total duration of the activity.  hh:mm:ss |
| **AverageHeartRate** | The average heart rate during the activity. |
| **MaximumHeartRate** | The maximum heart rate during the activity. |
| **AverageSpeed** | The average speed during the activity. Units must be either km/h or min/km |
| **MaximumSpeed** | The maximum speed during the activity. Units must be either km/h or min/km |
| **TotalAscent** | The total ascent during the activity. Units must be meters. |
| **TotalDescent** | The total descent during the activity. Units must be meters. |
| **AveragePower** | The average power output during the activity. |
| **Difficulty** | The difficulty level of the activity. |
| **MinimumTemperature** | The minimum temperature during the activity. Units must be celsius. |
| **MaximumTemperature** | The maximum temperature during the activity. Units must be celsius. |
| **MovingTime** | The duration of the activity while in motion.  hh:mm:ss |
| **ElapsedTime** | The total duration of the activity, including breaks.  hh:mm:ss |
| **MinimumAltitude** | The minimum altitude during the activity. Units must be meters. |
| **MaximumAltitude** | The maximum altitude during the activity. Units must be meters. |
| AverageRunningCadence | The average cadence during running activities |
| Maximum Running Cadence | The maximum cadence during running activities |
| Average Stride Length | The average stride length during running activities |
| Average Vertical Ratio | The average vertical ratio during running activities |
| Average Vertical Oscillation | The average vertical oscillation during running activities |
| Average Ground Contact Time | The average ground contact time during running activities |
| Average Pedaling Cadence | The average pedaling cadence during cycling activities |
| Maximum Pedaling Cadence | The maximum pedaling cadence during cycling activities |
| Training Stress Score | The training stress score for the activity |
| Maximum Power | The maximum power output during the activity |
| Fluency | The fluency of swimming strokes during swimming activities |
| Average Swolf | The average swolf score during swimming activities |
| Average Stroke Speed | The average stroke speed during swimming activities |
| Total Repetition | The total number of repetitions during strength training activities |
| Dive Time | The time spent diving during diving activities |
| Surface Interval | The time spent on the surface between dives |
| Normalized Power NP | The normalized power output during cycling activities |
| Left Right Balance | The left-right balance during cycling activities |
| Maximum Average Power 20 Minutes | The maximum average power output for 20 minutes during cycling activities |
| Total Strokes | The total number of strokes during swimming activities |
| Total Series | The total number of series during strength training activities |
| Average Rate Of Respiration | The average rate of respiration during the activity |
| Minimum Rate Of Respiration | The minimum rate of respiration during the activity |
| Maximum Breathing Rate | The maximum breathing rate during the activity |
| Change In Stress | The change in stress |

**Data cleaning:** The data is preprocessed in order to remove any incomplete or erroneous records, correct some typos and incoherencies, and remove useless or incomplete columns.

The typos that will be fixed are the following:

* Strings or values exported within quotes. The quotes will be removed.
* Numerical values exported as string will be turned into floats.
* Dates exported as string will be turned into datetime values.
* Activity types written in Spanish will be translated to English.
* Useless and uninformative columns will be removed.

**Contact information**: For any inquiries or questions about the dataset, please contact the creator at carlamiquelblasco2001@gmail.com