HBN Instruments

The tabular data in **train.csv** and **test.csv** comprises measurements from a variety of instruments. The fields within each instrument are described in **data_dictionary.csv**. These instruments are:

- Demographics - Information about age and sex of participants.

- Internet Use - Number of hours of using computer/internet per day.

- Children's Global Assessment Scale - Numeric scale used by mental health clinicians to rate the general functioning of youths under the age of 18.

- Physical Measures - Collection of blood pressure, heart rate, height, weight and waist, and hip measurements.

- FitnessGram Vitals and Treadmill - Measurements of cardiovascular fitness assessed using the NHANES treadmill protocol.

- FitnessGram Child - Health related physical fitness assessment measuring five different parameters including aerobic capacity, muscular strength, muscular endurance, flexibility, and body composition.

- Bio-electric Impedance Analysis - Measure of key body composition elements, including BMI, fat, muscle, and water content.

- Physical Activity Questionnaire - Information about children's participation in vigorous activities over the last 7 days.

- Sleep Disturbance Scale - Scale to categorize sleep disorders in children.

- Actigraphy - Objective measure of ecological physical activity through a research-grade biotracker.

- Parent-Child Internet Addiction Test - 20-item scale that measures characteristics and behaviors associated with compulsive use of the Internet including compulsivity, escapism, and dependency.

Note in particular the field PCIAT-PCIAT_Total. The target sii for this competition is derived from this field as described in the data dictionary: 0 for None, 1 for Mild, 2 for Moderate, and 3 for Severe. Additionally, each participant has been assigned a unique identifier id.

Actigraphy Files and Field Descriptions

During their participation in the HBN study, some participants were given an accelerometer to wear for up to 30 days continually while at home and going about their regular daily lives.

- **series_{train|test}.parquet/id={id}** - Series to be used as training data, partitioned by id. Each series is a continuous recording of accelerometer data for a single subject spanning many days.

    - id - The patient identifier corresponding to the id field in **train/test.csv**.

    - step - An integer timestep for each observation within a series.

    - X, Y, Z - Measure of acceleration, in *g*, experienced by the wrist-worn watch along each standard axis.

    - enmo - As calculated and described by the wristpy package, ENMO is the Euclidean Norm Minus One of all accelerometer signals (along each of the x-, y-, and z-axis, measured in g-force) with negative values rounded to zero. Zero values are indicative of periods of no motion. While no standard measure of acceleration exists in this space, this is one of the several commonly computed features.

    - anglez - As calculated and described by the wristpy package, Angle-Z is a metric derived from individual accelerometer components and refers to the angle of the arm relative to the horizontal plane.

    - non-wear_flag - A flag (0: watch is being worn, 1: the watch is not worn) to help determine periods when the watch has been removed, based on the GGIR definition, which uses the standard deviation and range of the accelerometer data.

    - light - Measure of ambient light in *lux*. See here for details.

    - battery_voltage - A measure of the battery voltage in *mV*.

    - time_of_day - Time of day representing the start of a 5s window that the data has been sampled over, with format %H:%M:%S.%9f.

    - weekday - The day of the week, coded as an integer with 1 being Monday and 7 being Sunday.

    - quarter - The quarter of the year, an integer from 1 to 4.

    - relative_date_PCIAT - The number of days (integer) since the PCIAT test was administered (negative days indicate that the actigraphy data has been collected before the test was administered).