Welcome to **Datathon of Technovation 2.0!**

Thank you for joining us at North South University's first-ever Datathon! This guide is meant to provide information about the different datasets you can work with and help you get off to a great start.

Getting Started

The target dataset to work on is given in Devpost Website. Your goal is to perform own analysis, extract useful insights. Examples of effective analysis include rigorous data visualization, hypothesis testing to solve a prediction problem, and building regression/classification models to learn something interesting. We have included possible methods of analysis for the datasets, but feel free to come up with your own!

Dataset Description

Our data-set for the contest contains the following information:

Raw accelerometer sensor data is collected from the smartphone and smartwatch at a rate of 20Hz. It is collected from 29 test subjects as they perform 6 activities (walking, jogging, ascending stairs, descending stairs, sitting, and standing) over a span of time.

Attribute Information

subject-id: ID describing the participant

activity-code: timestamp: Unix time (integer)

x: represents the sensor reading (accelerometer) for the x dimension

y: represents the sensor reading (accelerometer) for the y dimension

z: represents the sensor reading (accelerometer) for the z dimension

Methods of analysis:

- Find an appropriate learning algorithm/model and use it on this data-set to extract meaningful patterns and validate your model
- From the data, can you predict the portion of subjects healthier than others? Present your findings through any form of visualizations and validate your model.
- Analyze the common sequence of activities done by the subjects in the dataset within the time space
- Analyze the effect of three-axis values of accelerometer on the order of intensity of activities

HELP & COMMUNICATION

Make sure to join our Facebook page(https://www.facebook.com/technovationnsuacmsc/) for the event.

Judging Information

1) **Impact**: How important/interesting is the finding being found?
2) **Data Analysis and Visualization**: Are the findings presented in an aesthetically pleasing and easy to understand manner? Is the data analysis rigorous and mathematically sound?
3) **Creativity & Innovation** Are there advanced, creative modelling and analysis techniques being used? Are there non-orthodox analysis tools being used?
4) **Presentation** How clearly are the findings presented? Have the ideas been articulated well?
5) **Vision** Future plans for the project? Is there a clear idea of how they would like to use their findings?