Darknet Hidden Messages

# Regression Problem, 100pts

There are many things going on the darknet, positive and negative. During the last months, hidden messages have been captured by several the investigation groups. The messages seem to be generated from the correlation of several signals.

As a data scientist, you should identify the minimum set of signals (variables) to regenerate the hidden signal.

OBJECTIVE:

Define an algorithm to correctly recover the hidden signal. Use a limited number of variables for reconstruction.

FEATURES:

There are 10,000 features without descriptions.

TRAIN SET:

10,000 signals have been captured that include the hidden message (objective).

TEST SET:

There are 3374 observations for which the hidden signal needs reconstruction.

LINK TO DATA:

Train and Test sets can be found in this folder.

DELIVERABLES:

* Summary (must): Explain your approach to solve the problem to any hackathon supporter of the Data Science challenge.
* Regression: Share your regression results of the test set. Include only the ID, SIGNAL\_VALUE columns. You will get back your R^2 score\*

**\*** Your R^2 score represents the number of points you will obtain. Zero being 0 points and One being 100 points. You can share your regression as many times as you wish, your highest value will be reported as the number of points obtained.