# Challenge:

To build a regression model that predicts the diameter of cylinder (a scalar value, Y) from a large number of impedance measurements (a 1568-dimensional vector, X).

# Description of data:

## Training data:

There are 84007 training samples.

The file *train* contains a 84007x1569 matrix, in CSV format. For each row, the first 1568 columns are the 1568 dimensions of the input vector, X. The last column is the output scalar, Y.

## Test data:

There are 5 separate sets of testing data, contained in files *test1, test2,… test5*.

The difficulty of the training set increases with the index. So, *test2* is more difficult than *test1*, but is easier than *test3*.

Note: DO NOT use any test data for training.

Each file *testX* contains a 21x1569 matrix, in CSV format. For each row, the first 1568 columns are the 1568 dimensions of the input vector, X. The last column is the output scalar, Y.