A note on health index

Introduction

Health index is introduced to allow fine grained control on service or equipment abnormalities. Its value is limited within [0,100], 0 being worst and 100 best. Its value is calculated for each prediction. Any downward trend or sudden dip will act as early warning and may initiate investigation.

Concept

Health index is dependent on factor aggregated likelihoods that prediction algorithm provides and its relationship with actual failure. Individual factor (feature) likelihood is calculated by predicting failure w.r.t. to each factor using proprietary superposed sequence.

Ideally, health index should be around 100 when factor aggregated likelihood values reaches lowest and around 0 when factor aggregated likelihood value reaches highest. In the intervening period, we should have a continuous function with following properties:

1. It should reach very close to zero when we reach average of the values corresponding to failure cases.
2. It should asymptotically reach zero with decreasing value of factor aggregated likelihood
3. It should reach very close to hundred when we reach lowest value of factor aggregated likelihood
4. It should asymptotically reach 100 with decreasing value of factor aggregated likelihood

Calculation

We use a modified sigmoid curve so that it stays between 0 and 100 over the range of only positive values. Sigmoid curve is used since its continuous and also provides asymptotic behaviour on both sides. The exact equation is provided below:

$f(s,x,y) = (1 - (2*(1/(1+EXP(-1*IF(s<x,0,s-x)/y)))-1))*100$

s – Factor aggregated likelihood
x - Offset
y - Scaling factor

Parameters x and y are estimated as follows:

Offset estimation:

Find the maximum of non-failure factor aggregated likelihoods. This means that if the factor aggregated likelihoods remain below this value, we can assume that tool is in healthy condition. Its value has to be decided after checking multiple tool predictions and it is safe to assume it little lower when in doubt.

Scaling factor estimation:

a. Find out average of factor aggregated likelihoods during failure
b. Subtract offset from the value found in step a

    c.   Divide the value found at step b by 4 (Sigmoid curve reaches almost 1 when x-variable is 4)
    d.   Use the value found at step c as Scaling factor

Thresholds:

Since health index is calculated based on data characteristics of each tool type, the threshold values remain invariant across tools. We propose following thresholds:

    a.   Failure: Less than or equal to 20
    b.   Warning: Greater than 20 and less than or equal to 40. Please note that health index changes smoothly from lower to higher value after a PM activity is undertaken. So, warning should always be considered whether we are recovering after a PM or a real warning has happened.
    c.   Fine: Greater than 40

Cumulative Health Index (CHI)

If HI is above threshold i.e. greater than 40, then HI is kept unchanged. Otherwise, we check the previous four HI data points, consider the three minimums (including current one), and take the average of these. If the previous four data points are not available, we consider as many as available. Please note that CHI is always based on health index and not on previous CHI.