Predicting Gas-Refill Time

# Introduction

Attached data belongs to a machine which runs on a gas, known as c-gas. When the c-gas, in gas chamber of machine, depletes, or is replaced by exhaust gases, the machine stops. The chamber has to be opened up to let the exhaust gases flush away and c-gas is filled up again.

It is difficult to calculate the amount of c-gas in chamber but proportion of the c-gas can be calculated. When this proportion reaches a minima the machine is stopped and refilled.

Machine beeps when the c-gas is at lowest and a manual operator refills it. C-gas is difficult to store and hence transporting it to machine takes lot of time and effort. If the refill time of the c-gas can be predicted before-hand, then it is possible to plan the transport and refill it when machine is still running.

Duration from the time of refill till the stopping point of machine, is termed as a Cycle.

The time is calibrated to start at 100 time units (which is the zero time). From then the proportion of the gas is calculated at each time-unit. If the prediction is made before 1800 time units then the transport and refill can be managed before depletion. But the sooner the better.

(Time unit is the custom time unit for machine. Each machine time unit is approx.. 45 minutes. But no conversion is necessary for this problem as the prediction needs to be in same units.)

# Problem Statement

Using the proportion of the gas in the chamber, the refill time (End of cycle time) needs to be determined before 1800 time units.

# Data:

Training data consists of 40 cycles. Data per time unit is provided for each cycle in long format. The zero time starts from 101 for all cycles.

|  |  |  |  |  |
| --- | --- | --- | --- | --- |
| **Cycle** | **Type** | **EndTime** | **Time** | **Values** |
| 2 | 2 | 3490 | 101 | 20.402 |
| 2 | 2 | 3490 | 102 | 20.402 |
| 2 | 2 | 3490 | 103 | 20.402 |
| 2 | 2 | 3490 | 104 | 20.402 |
| 2 | 2 | 3490 | 105 | 20.402 |
| 2 | 2 | 3490 | 106 | 20.383 |
| 2 | 2 | 3490 | 107 | 20.383 |
| 2 | 2 | 3490 | 108 | 20.383 |
| 2 | 2 | 3490 | 109 | 20.383 |
| 2 | 2 | 3490 | 110 | 20.383 |
| 2 | 2 | 3490 | 111 | 20.36 |
| 2 | 2 | 3490 | 112 | 20.36 |
| 2 | 2 | 3490 | 113 | 20.36 |
| 2 | 2 | 3490 | 114 | 20.36 |
| 2 | 2 | 3490 | 115 | 20.36 |
| 2 | 2 | 3490 | 116 | 20.36 |
| 2 | 2 | 3490 | 117 | 20.36 |

**Cycle**: Duration from the time of refill, till the end of that refill is called a cycle. The column cycle is just a serial number to represent in a long format. (Redundant feature used only for naming the cycle)

**Type**: Type no. is a type of machine. The solution needs to be invariant of type hence it’s a categorical feature of the machine.

**Values**: C-gas proportion measured per time unit.

**Time**: time units passed

**End** **Time**: The time at which the cycle ends. Can also be determined by checking the time value in the end of the cycle in training data.

**Testing:**

The test data will only have cycles till time 1800 time units. End time needs to be predicted for each cycle. Sample submission is attached.

# Evaluation criteria

Mean absolute error