**Workshop 1(a)**

**Proposed Architecture:**

Linear Regressor

Excel Input

GA optimization

Estimated User

Clicks

Constraints

Initial

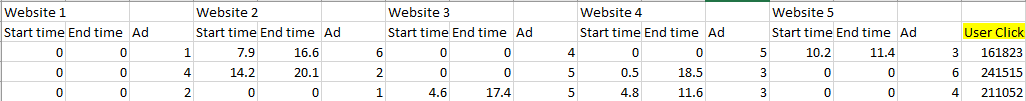
Chromosome

Objective

2-Step Approach details:

1)- Building a Linear Regression Model:

Train a regression model using this 1000-record dataset, using the 15 input variables and actual number of user clicks.



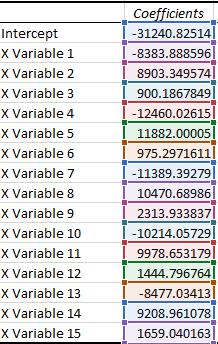
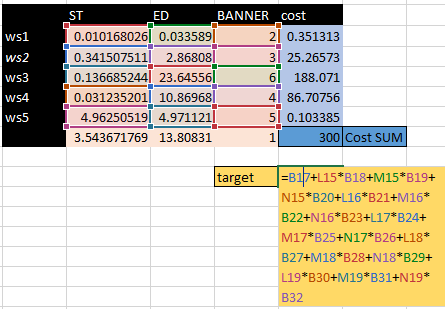
2)- Use GA to find the optimal solution for the 15 input variables.

Once we had Coefficient for Regression that approximates the relationship between the actual inputs and the actual user clicks in the past 3 years.

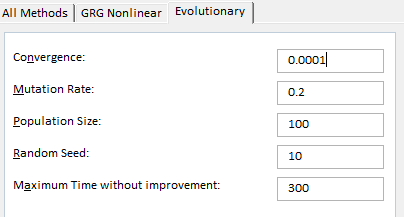
We use this Coefficient for our Target function to maximize which represent the Number of clicks based on 15 input variables.

*f'*(*X*) = *β*0 + *β*1*X*1 + *β*2*X*2 + *. . .* + *β15X15*

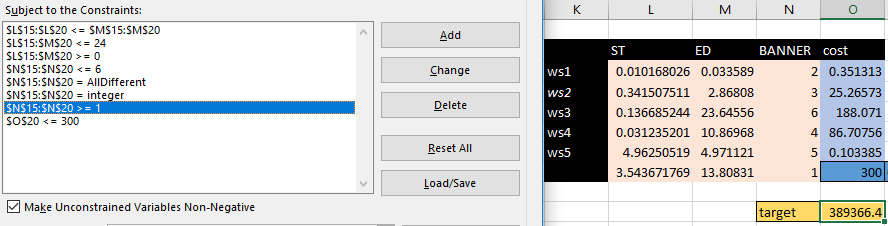
Where *β0 ,…, β15* is Coefficient from regression and X1,….,X15 are input variables which will be optimized(Gene).



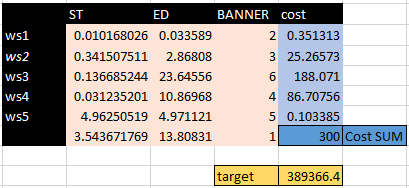
We try different GA settings to get better result, our final settings showed below:



Below are the Constraints used for GA optimization:



Below is the Maximized objective with given constraints, provided setting and Regression model used:



**Workshop 1(b)**

In this workshop, instead of using a regression model to approximate the relationship between the actual recorded inputs and the actual number of user clicks.

1B suggests using an alternative model based on ‘expert experience’ to model this relationship:

**heuristic equation** = user clicks achievable from adj for display duration D using website

W= (dur-1\*clk-1+dur-2\*clk-2+dur-3\*clk-3) \* scalefactor*j*

Where clk-i is the user clicks achievable per hour for the i-th time period (i = 1 to 3) and dur-i is the time duration that adj (j = 1 to 6) is displayed on website W during the i-th time period.

**[Note** dur-1, dur-2, dur-3 form a continuous time period and dur-1 + dur-2 + dur-3 = D]

Unlike previous model where we can find the parameters of the model using regression, here we do not know values of parameter of our ‘*heuristic’* model, so in our case we need to find the given parameters as shown below:

Details Parameter

We split a 24-hour day into 3 parts using cp1 and cp2 - cp1, cp2

Click in 3 duration with cp1 and cp2 for each of the 5 websites - clk1, clk2, clk3

For banners 1 to 6 - scalefactor

Visualization:

O hr

24 hr

CP1

CP2

Dur1

Dur2

Add ST

CLK1

CLK2

CLK3

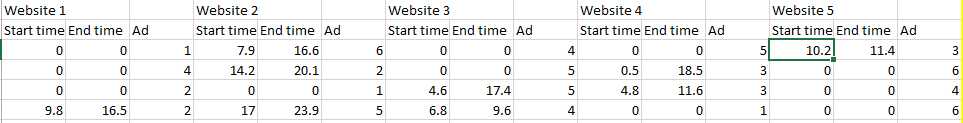
Add ET

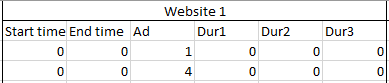
The start time and end time for the displayed ad can fall into any of these 6 scenarios -

1. ST | ET | CP1 | CP2
2. ST | CP1 | ET | CP2
3. CP1 | ST | CP2 | ET
4. CP1 | CP2 | ST | ET
5. ST | CP1 | CP2 | ET
6. CP1 | ST | ET | CP2

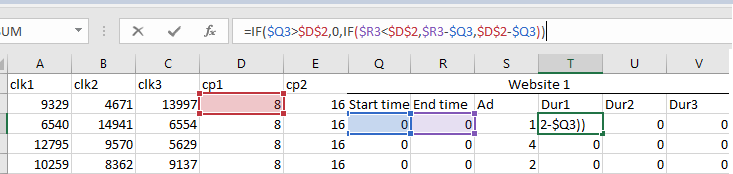
**Data Preparation:**

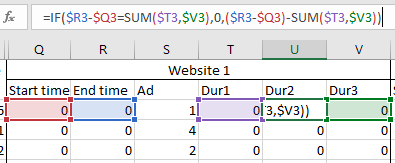
**Raw Data:**

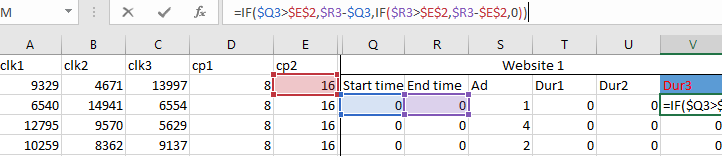


Calculate the Dur1,Dur2 and Dur3 between CP1 and CP2 for every Websites. Like

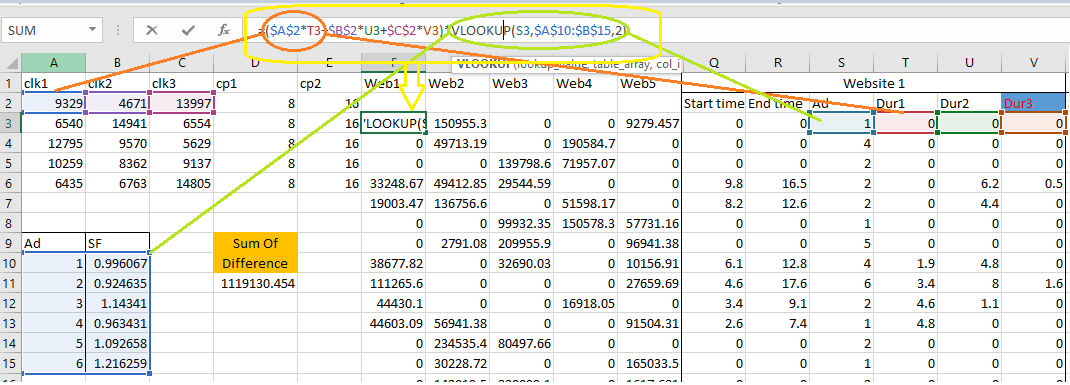
Calculation of Dur1 , Dur2 and Dur3 are below respectively:





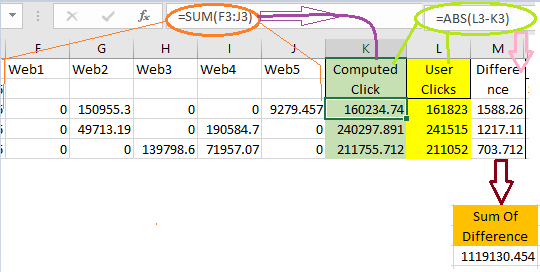


We have also calculated click for every website using the ‘Heuristic’ formula given above, by addition of multiplication between Duration(1,2,3) and Click(1,2,3) respectively and Multiplying whole unit by Scale Factor(1,2,3,4,5,6) for every ad banners.

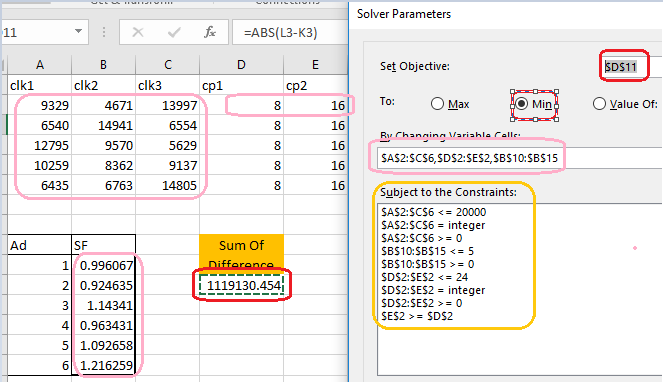


In Next step we will Do Sum for all individual website computed click and termed it as “Computed Click”.

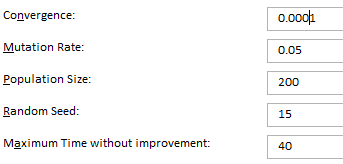
We already have for all population Original click done by user, So what we planned is to do comparison between Actual value vs. Predicted value and try to **Minimize this value as** **Objective function** in GA.



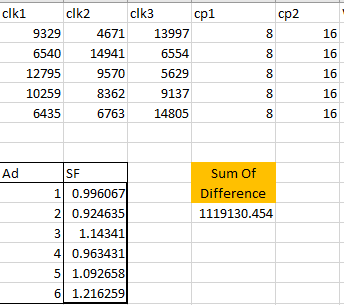
Solver Setting up details:



Solver Options:

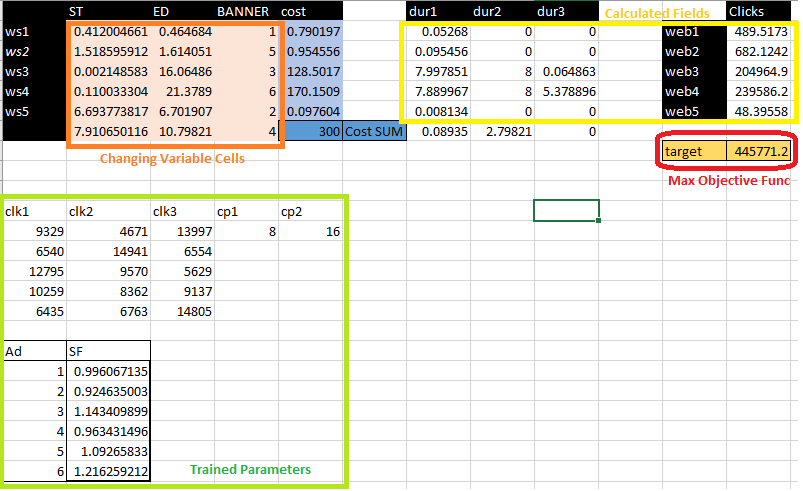


**Final result** for parameter Optimization done using Evolutionary GA:



Now that we have the values of the optimized parameters of our heuristic model, we can discard the previously used regression .Next, with our GA-optimized heuristic model, we are ready to run GA to find the best input combination that will give the highest user clicks.

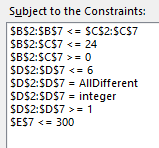
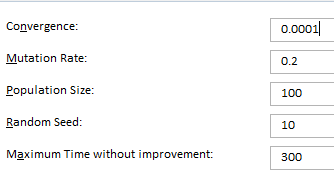
**Running GA to find the best input combination using Heuristic parameters GA.**



In above Screenshot, we have 4 parts:

1. Value to be optimized (In orange cell)
2. Trained Parameters , previously trained in 1B part(In green cell)
3. Calculated columns for each website clicks using Parameters and *‘Heuristic’* formula (In yellow cell)
4. Target Objective function which need to be maximized.

We have used GA setting as below:



We are getting Clicks of : **44571** clicks with all constraints.



Testing this plan with the OnlineAds.exe application, we get 470526.

Comparing with Regression model and ‘Heuristic’ model we achieve better score with Heuristic:

