A Study of Eye Tracking Data based on Multiple Regression Analysis

Gufran Ahmad
Global IT Consultancy for Innovations, India
gufran.researcher@gmail.com

Abstract

Exploration of eye tracking data from a number of perspectives can enlighten the human cognition as well as human interaction. In this eye tracking study, we intended to bring potentiality of data insights through prominent and effective regression analysis techniques. At first, we set up experimentations and collected eye tracking data of participants. Next, we preprocessed the data for precision and optimization of data products in terms of resultant outcomes. Finally, the statistical techniques of regression analysis were engaged in the context analytically and resultant data were interpreted for comprehension of visual perception.

Keywords: *Eye tracking, Regression analysis, Cognition, Visual perception*

1. Introduction and background

As the eye tracking has leaped and scaled for better understanding of visual perception and the dynamics of human mind, an extensive range of methodologies and treatments are exercised and developed for attempting numerous and unknown problems in this area of eye tracking researches. Further, eye tracking has become one of the sophisticated and significant research studies of this modern world. In recent times, commercial eye tracking products and advanced researches are being acquainted with the latest tools and techniques of data science and machine learning to improve the quality of data as well as to predict better the unknown problems with a view to cognitive comprehension [1-5].

In the eye tracking experimentations, human eyes shift their gazes and visual focus moves for intended specific purposes to retrieve information. The fixations and saccades are two basic activities within every eye movements. The tendency for information seeking and gaining visual attention and visual perception are the processes that are continuously dominating along with other cognitive and metacognitive processes, like flow of analogical thoughts, associative relevancy and cognitive stimulus, decision-making processes, etc. The movements of eyes are recorded by eye tracking systems. Such recorded tracking of eye movements gives crucial insights about the undergoing and underlying processes that get involved during scene viewing [6-23].

Moreover, these underlying processes generate crucial interrelations that create resultant patterns of eye fixations. Therefore, it is essential to estimate and measure the existing relationships and parametric dependency and variability within these generated and collected data of eye tracking. Further, it is the measurement of eye activities in terms of human cognition, especially visual perception that can be applied to visualization, eye tracking analysis as well as iris detection. [6-23].

Multiple regression analysis is the study of establishing relationships among variables and the multiple regression techniques are employed in data for better insights of relevant outcomes. In other word, multiple regression analysis can be considered as a method of measuring the relations among two or more facts. Multiple regression analysis includes many techniques for modeling and analyzing numerous variables. These regression techniques are the best fit for the scenarios when the emphasis is on the relationships among a dependent variable and one or more independent variables. These independent variables are also called as predictors or explanatory variables. The dependent variables are also known as criterion variables [24-37].

Specifically, multiple regression analysis facilitates to identify how the distinctive value of the dependent variable varies when a set of independent variables are modified. Here, we are interested in investigating more than one independent variable of our dependent variable. The motive behind these arrangements is to find out whether the presence of added independent variables directs towards augmented expectation of the resultant variable. The techniques and analysis of regression have become de facto in providing solutions to the complex problems of numerous sciences and interdisciplinary sciences [24-37].

2. Present study plan

By the current research study, we investigate the eye tracking data collected from participants' eye tracking. The data is analyzed quantitatively using the multiple regression analysis techniques. The multiple regression analysis approaches that are engaged on the eye tracking data can illustrate and elucidate the characteristics and relationships among various variables of eye tracking data and their interdependencies.

With the tools of multiple regression analysis, we detect and estimate the best-fit model of multiple regression among various variables of eye tracking data. These measurements are essential part of this research work. By analyzing the data, we can explore the data insights that are significant for establishing the inherent mechanism of cognitive processes and the dynamics of visual perception. Therefore, the modeling of multiple regression predicts the dependencies of processes associated with visual attention and visual perception.

The multiple regression analysis produces a number of data insights statistically that are interpreted and elaborated in accordance with the underlying mechanism and processes of visual attention and visual perception. The analyzed outcomes and existing relationships among the variables in terms of numeric quantities and statistical indicators are resilient facts that reflect the undergoing phenomena of human cognition and visual perceptions. By mentioning these values and interpreting their meanings through data interpretations, we can establish a consolidated pitch for our research findings.

In other words, we examine how multiple independent variables or predictors can be used to model a dependent variable or response variable. This approach is a generalized and extended method of simple linear regression model that are used to study of statistically significant relationship between two variables. Here, in our study, we have a set of complex variables of eye tracking data. Hence, we use this multiple regression analysis and modeling techniques which are the most suitable and exact approach to handle the current research problem.

In the most simplified form and linear representation, we can write the multiple regression model as below.

$$Y = \beta_0 + \beta_1 X_1 + \beta_2 X_2 + \dots\dots\dots\dots + \beta_p X_p + \varepsilon$$

where, the dependent variable or response variable is denoted as Y, the number of independent variables is represented as p, the independent variables are indicated as ($X_1, X_2, \dots, X_p$), coefficients (parameters) are denoted as ($\beta_1, \beta_2, \dots, \beta_p$), the error term is depicted as $\boldsymbol{\varepsilon}$ which is $N(0, \sigma^2)$, and σ is known as the standard deviation of the population.

The schematic diagram of eye tracking system and present research study plan as process that involved during eye tracking experimentation are represented in Figure 1.

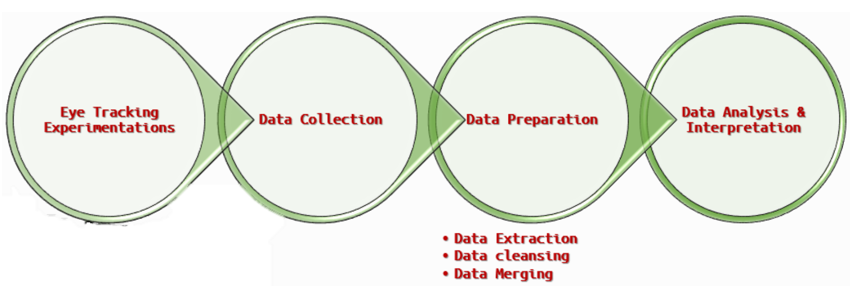

Figure 1. Present research study process.

3. Experimental setup and procedure

At first, we selected an artistic scenery, Green Hills, for our experimentation. The reason to choose this artistic scenery was having numerous cognitively built in human emotions and feelings that were our utmost agenda to trace or record these cognitively generated human activities during eye tracking of the Subjects, i.e. participants.

The experimental setup consisted of eye tracking system that was used for recording of eye tracking. In eye tracking system, the system illuminated infrared light for tracking the eye tracking. The camera, connected to the system, captured the location of viewer's eyes in terms of tracking during experimentation time. As the viewer moved his/her eyes to look a new location of the scene, the camera recorded new tracking also. This process of recording continued subsequently. The system-generated eye tracking trails and heat maps using the captured data that was utilized for further analysis.

The schematic diagram of eye tracking system and basic processes involved during eye tracking experimentation was represented in Figure 2.

Figure 2. Operational processes of eye tracking system with Green Hill scenery.

In our experiments, we studied track of eye tracking as the sequenced gazing of viewer's eye tracking, which was generated by the system, during scene viewing. These were the dynamic shifts of eye gaze in scene viewing. By these eye fixations, tracking pattern was generated by eye tracking system that records the human eye tracking.

4. Method and data collection

We selected 40 participants from a number of fields randomly, aging from 19 years to 45 years. Further, we assigned these participants (Subjects) to view selected famous artistic scenery (Object).

Their eye tracking were closely monitored as they viewed 32 bits full-color artistic scenes. The Objects, the scenes were displayed on a computer monitor. We have shown the scenery at a resolution of 1280 × 1024 pixels and subtended 15 degree horizontally by 10 degree vertically at a viewing distance of 75 centimeter. Eye position was sampled from an Eye Tech Digital Systems TM3 16 mm Eye Tracker, and tracking data was parsed into sequenced gazing with circles of concentration.

The Subjects' heads were held steady in advance prior to experimentation. Prior to the first trial, Subjects completed a procedure to calibrate the output of the eye tracker against spatial positions on the display screen. This procedure was repeated regularly throughout the experiment to maintain high level of accuracy. Subjects were initiated to view the scenes freely.

The scenery was presented to the Subjects in very comfortable mode. During the time span, the Subjects viewed the scenes with their normal eyes and focused attention on the Object, i.e. the scenery.

After the phase of data collection, we started the process of data preparation that involved tediously longest phase of this study. This process of data preparation consisted of data extraction, data cleansing, and data merging. At first, the data was extracted and transferred from the various sources of the eye tracking system. Next, a mapped transformation of data was loaded by the data analysis system. In our case, we utilized the data analysis system equipped with R environment. In the next step, the eye tracking data were cleaned for all kinds of ambiguity, defect, and inconsistency by using a set of approaches to refine and tidy the eye tracking data. Finally, the data were combined and merged as a complete set of all samples to go ahead for data analysis phase.

5. Data analysis and data interpretation

During this phase of analysis, we analyzed statistically the eye tracking data for estimation and evaluation of existing variables of the data population in terms of numerous data samples.

Although we conducted and carried out a number of statistical analysis for this eye tracking data, yet we presented those results that seemed to be appealing for conclusive remarks and did viable evidences within our statistical population.

We computed and analyzed all of our data generated by eye tracking system using R environment.

We started with the computation of the statistical summary of the collected eye tracking data that turned out to be as shown in Table 1.

The variables of eye tracking data intended such as,

1. the tracking time in millisecond(Time[msec]),
2. ticking time during eye tracking(Time[Ticks]),
3. gaze X and Y coordinates (GazeX, GazeY),
4. the diameter of eye tracking focused circle(Diameter),
5. the left and right calibrations of the eye tracking device in terms of left and right eyes
6. (LCalib, RCalib)
7. the left and right cross sectional positions in terms of left and right eyes (LFound, RFound)
8. the left eye's X coordinate, Y coordinate, and diameter of focused circle (LX, LY, LD)
9. the right eye's X coordinate, Y coordinate, and diameter of focused circle (RX, RY, RD)
10. the positional accuracy of eye gaze in terms of logical; FALSE or TRUE (Lost), can be considered as categorical variable

Table 1. The statistical summary of the collected eye tracking data

Time[msec]	Time[Ticks]	GazeX	GazeY	Diameter	LCalib	LFound	LX
Min. : 15.62	Min. :6.342e+17	Min. : 0.0	Min. : 3.0	Min. :1.600	Min. :1	Min. :1	Min. :-336.0
1st Qu.: 6000.00	1st Qu.:6.342e+17	1st Qu.: 344.0	1st Qu.:465.0	1st Qu.:2.500	1st Qu.:1	1st Qu.:1	1st Qu.: 347.0
Median : 12484.38	Median :6.343e+17	Median : 654.0	Median :568.0	Median :2.800	Median :1	Median :1	Median : 655.0
Mean : 20409.14	Mean :6.342e+17	Mean : 636.8	Mean :574.6	Mean :2.767	Mean :1	Mean :1	Mean : 638.8
3rd Qu.: 22234.38	3rd Qu.:6.343e+17	3rd Qu.: 887.0	3rd Qu.:676.0	3rd Qu.:2.800	3rd Qu.:1	3rd Qu.:1	3rd Qu.: 886.0
Max. :101390.62	Max. :6.343e+17	Max. :1279.0	Max. :959.0	Max. :4.900	Max. :1	Max. :1	Max. :1716.0

LY	LD	RCalib	RFound	RX	RY	RD	Lost
Min. : -14.0	Min. :1.60	Min. :1	Min. :1	Min. : 5.0	Min. : 21	Min. :1.600	Mode :logical
1st Qu.: 466.0	1st Qu.:2.20	1st Qu.:1	1st Qu.:1	1st Qu.: 345.0	1st Qu.: 460	1st Qu.:2.200	FALSE:18807
Median : 569.0	Median :2.80	Median :1	Median :1	Median : 652.0	Median : 566	Median :2.800	
Mean : 584.8	Mean :2.74	Mean :1	Mean :1	Mean : 638.1	Mean : 581	Mean :2.794	
3rd Qu.: 678.0	3rd Qu.:2.80	3rd Qu.:1	3rd Qu.:1	3rd Qu.: 888.0	3rd Qu.: 677	3rd Qu.:2.800	
Max. :2272.0	Max. :5.20	Max. :1	Max. :1	Max. :2217.0	Max. :2685	Max. :4.600	

In this statistical summary of the collected eye tracking data, the data population for this eye tracking data indicated normal tendency and all the variables of eye tracking data had no abnormality. Further, the variable, Time [Ticks], had exponential factor in its measurements. Moreover, all the variables of eye tracking data had numeric data quantities except for the variable, Lost, that had nonnumeric data quantities, Boolean or logical data.

Next, we started the process of multiple regression analysis by analyzing the data for all variables sequentially. Each analysis involved with two and more variables during the process of modeling multiple regression.

The significant outcomes that were of prime interest to us had shown as below sequentially.

Table 2. Multiple regression analysis for dependent variable, Time [msec]

```
Call:
lm(formula = eye.data$`Time[msec]` ~ `Time[Ticks]` + GazeX +
    Diameter + LX + LY + LD + RX + RY, data = eye.data)

Residuals:
   Min     1Q Median     3Q    Max
-44073  -7682  -1367   6914  59864

Coefficients:
                Estimate Std. Error  t value Pr(>|t|)
(Intercept)    2.555e+08  2.299e+06  111.139  < 2e-16 ***
`Time[Ticks]` -4.027e-10  3.623e-12 -111.146  < 2e-16 ***
GazeX          1.318e+01  7.212e+00    1.828  0.06755 .
Diameter      -1.521e+04  9.938e+02  -15.305  < 2e-16 ***
LX            -1.654e+01  5.225e+00   -3.165  0.00155 **
LY             6.287e+00  1.992e+00    3.156  0.00160 **
LD             8.469e+03  8.865e+02    9.553  < 2e-16 ***
RX             5.937e+00  3.209e+00    1.850  0.06427 .
RY             5.676e+00  2.027e+00    2.801  0.00510 **
---
Signif. codes:  0 ‘***’ 0.001 ‘**’ 0.01 ‘*’ 0.05 ‘.’ 0.1 ‘ ’ 1

Residual standard error: 16710 on 18798 degrees of freedom
Multiple R-squared:  0.4694,    Adjusted R-squared:  0.4692
F-statistic:  2079 on 8 and 18798 DF,  p-value: < 2.2e-16
```

As we observed from Table 2, the multiple regression model based on dependent variable, Time [msec] significantly depended on Time [Ticks], Diameter, LX, LY, LD, and RY. The probabilities of these independent variables (in the fifth column of the table) came out to be much lower than 0.05 (i.e. the 95% confidence level or more). Further, the estimated values of coefficients, (β1, β2, …, βp) for independent variables, Time [Ticks], Diameter, LX, LY, LD, and RY were computed in second column of the table. The values of R-squared and adjusted R-squared were found to be of good-sized. Additionally, residual standard error and F values were also suggestively high numbers.

Table 3. Multiple regression analysis for dependent variable, Time [Ticks]

```
Call:
lm(formula = eye.data$`Time[Ticks]` ~ `Time[msec]` + GazeX +
    GazeY + Diameter + LX + LY + LD + RX + RY, data = eye.data)

Residuals:
       Min         1Q     Median         3Q        Max
-8.662e+13 -1.070e+13  4.001e+12  1.767e+13  6.267e+13

Coefficients:
              Estimate Std. Error    t value Pr(>|t|)
(Intercept)  6.343e+17  1.404e+12 451924.649  < 2e-16 ***
`Time[msec]` -9.796e+08  8.851e+06   -110.671  < 2e-16 ***
GazeX       -7.318e+10  1.151e+10     -6.358 2.09e-10 ***
GazeY       -3.829e+10  3.860e+09     -9.919  < 2e-16 ***
Diameter    -9.088e+13  1.415e+12    -64.209  < 2e-16 ***
LX           3.148e+10  8.307e+09      3.790 0.000151 ***
LY           6.710e+10  3.599e+09     18.646  < 2e-16 ***
LD           6.523e+13  1.303e+12     50.058  < 2e-16 ***
RX           3.339e+10  5.062e+09      6.595 4.35e-11 ***
RY          -3.445e+10  3.360e+09    -10.255  < 2e-16 ***
---
Signif. codes:  0 ‘***’ 0.001 ‘**’ 0.01 ‘*’ 0.05 ‘.’ 0.1 ‘ ’ 1

Residual standard error: 2.607e+13 on 18797 degrees of freedom
Multiple R-squared:  0.5831,    Adjusted R-squared:  0.5829
F-statistic:  2922 on 9 and 18797 DF,  p-value: < 2.2e-16
```

The multiple regression analysis over dependent variable, Time [Ticks] (in Table 3) showed important outcome, as the dependent variables, Time [msec], GazeX, GazeY, Diameter, LX, LY, LD, RX, and RY got significantly very low probabilities. This meant the dependency depicted by this multiple regression model had very high level of confidence. However, the residual standard error had increased drastically.

Table 4. Multiple regression analysis for dependent variable, Diameter

```
Call:
lm(formula = eye.data$Diameter ~ `Time[msec]` + `Time[Ticks]` +
    GazeX + GazeY + LD + RD, data = eye.data)

Residuals:
       Min         1Q     Median         3Q        Max
-7.400e-14 -5.000e-15 -2.000e-15  1.000e-15  3.415e-11

Coefficients:
                Estimate Std. Error    t value Pr(>|t|)
(Intercept)   -1.280e-11  4.370e-11 -2.930e-01   0.7696
`Time[msec]`  -3.315e-20  1.086e-19 -3.050e-01   0.7601
`Time[Ticks]`  2.058e-29  6.889e-29  2.990e-01   0.7652
GazeX         -9.885e-18  5.769e-18 -1.713e+00   0.0867 .
GazeY          1.018e-17  1.220e-17  8.340e-01   0.4042
LD             5.000e-01  6.429e-15  7.777e+13   <2e-16 ***
RD             5.000e-01  7.442e-15  6.719e+13   <2e-16 ***
---
Signif. codes:  0 ‘***’ 0.001 ‘**’ 0.01 ‘*’ 0.05 ‘.’ 0.1 ‘ ’ 1

Residual standard error: 2.491e-13 on 18800 degrees of freedom
Multiple R-squared:      1,     Adjusted R-squared:      1
F-statistic: 1.106e+28 on 6 and 18800 DF,  p-value: < 2.2e-16
```

The multiple regression analysis for the dependent variable, Diameter (Table 4) turned out to be significant only for two independent variables, LD, and RD. Rest of the variables had much higher probabilities than the standard 0.05 (95% confidence level). Additionally, F value was found to be high enough but residual standard error had very low quantity.

Table 5. Multiple regression analysis for dependent variable, GazeX

```
Call:
lm(formula = eye.data$GazeX ~ `Time[msec]` + `Time[Ticks]` +
    GazeY + Diameter + LX + LY + LD + RX + RY, data = eye.data)

Residuals:
     Min       1Q   Median       3Q      Max
-313.691   -3.238    0.314    4.279  197.384

Coefficients:
                Estimate Std. Error t value Pr(>|t|)
(Intercept)    1.863e+04  2.926e+03   6.366 1.99e-10 ***
`Time[msec]`   1.292e-05  7.201e-06   1.795   0.0727 .
`Time[Ticks]` -2.933e-14  4.613e-15  -6.358 2.09e-10 ***
GazeY          7.269e-02  2.392e-03  30.392  < 2e-16 ***
Diameter      -7.883e+00  9.877e-01  -7.981 1.53e-15 ***
LX             6.572e-01  2.168e-03 303.057  < 2e-16 ***
LY            -2.934e-02  2.289e-03 -12.818  < 2e-16 ***
LD            -1.742e+00  8.781e-01  -1.983   0.0473 *
RX             3.300e-01  2.121e-03 155.567  < 2e-16 ***
RY            -2.964e-02  2.122e-03 -13.969  < 2e-16 ***
---
Signif. codes:  0 ‘***’ 0.001 ‘**’ 0.01 ‘*’ 0.05 ‘.’ 0.1 ‘ ’ 1

Residual standard error: 16.5 on 18797 degrees of freedom
Multiple R-squared:  0.9976,    Adjusted R-squared:  0.9976
F-statistic: 8.714e+05 on 9 and 18797 DF,  p-value: < 2.2e-16
```

The multiple regression model for dependent variable, GazeX (Table 5) showed significantly lower probabilities for independent variables, Time [Ticks], GazeY, Diameter, LX, LY, LD, RX, and RY. The F value was found to be muchlarge.

Table 6. Multiple regression analysis for dependent variable, GazeY

```
Call:
lm(formula = eye.data$GazeY ~ `Time[Ticks]` + GazeX + Diameter +
    LX + LY + LD + RX + RY, data = eye.data)

Residuals:
    Min      1Q  Median      3Q     Max
-821.73  -17.17    4.18   22.49  138.05

Coefficients:
                Estimate Std. Error t value Pr(>|t|)
(Intercept)    8.608e+04  6.756e+03  12.740  < 2e-16 ***
`Time[Ticks]` -1.356e-13  1.065e-14 -12.732  < 2e-16 ***
GazeX          6.443e-01  2.120e-02  30.395  < 2e-16 ***
Diameter       2.611e+01  2.921e+00   8.938  < 2e-16 ***
LX            -4.183e-01  1.536e-02 -27.237  < 2e-16 ***
LY             4.858e-01  5.855e-03  82.969  < 2e-16 ***
LD            -1.505e+01  2.606e+00  -5.777 7.71e-09 ***
RX            -2.068e-01  9.432e-03 -21.925  < 2e-16 ***
RY             3.008e-01  5.957e-03  50.485  < 2e-16 ***
---
Signif. codes:  0 ‘***’ 0.001 ‘**’ 0.01 ‘*’ 0.05 ‘.’ 0.1 ‘ ’ 1

Residual standard error: 49.13 on 18798 degrees of freedom
Multiple R-squared:  0.9038,    Adjusted R-squared:  0.9038
F-statistic: 2.209e+04 on 8 and 18798 DF,  p-value: < 2.2e-16
```

The multiple regression analysis for dependent variable, GazeY (Table 6) indicated that the probabilities of dependent variables, Time [Ticks], GazeX, Diameter, LX, LY, LD, RX, and RY were very low, i.e. the levels of confidence were found to be very high. The residual standard error had relatively higher value and F value was higher than typical value as well.

Table 7. Multiple regression analysis for dependent variable, LX

```
Call:
lm(formula = eye.data$LX ~ `Time[msec]` + `Time[Ticks]` + GazeX +
    GazeY + Diameter + LY + LD + RX + RY, data = eye.data)

Residuals:
    Min      1Q  Median      3Q     Max
-322.09   -5.82   -0.51    4.98  402.30

Coefficients:
                Estimate Std. Error t value Pr(>|t|)
(Intercept)   -1.542e+04  4.059e+03  -3.799 0.000146 ***
`Time[msec]`  -3.107e-05  9.982e-06  -3.113 0.001854 **
`Time[Ticks]`  2.426e-14  6.400e-15   3.790 0.000151 ***
GazeX          1.263e+00  4.168e-03 303.057  < 2e-16 ***
GazeY         -9.072e-02  3.332e-03 -27.230  < 2e-16 ***
Diameter       8.093e+00  1.370e+00   5.905 3.58e-09 ***
LY             9.267e-03  3.187e-03   2.908 0.003645 **
LD             6.074e+00  1.217e+00   4.992 6.02e-07 ***
RX            -2.533e-01  4.047e-03 -62.581  < 2e-16 ***
RY             7.069e-02  2.912e-03  24.279  < 2e-16 ***
---
Signif. codes:  0 ‘***’ 0.001 ‘**’ 0.01 ‘*’ 0.05 ‘.’ 0.1 ‘ ’ 1

Residual standard error: 22.88 on 18797 degrees of freedom
Multiple R-squared:  0.9955,    Adjusted R-squared:  0.9955
F-statistic: 4.585e+05 on 9 and 18797 DF,  p-value: < 2.2e-16
```

The model of multiple regression for dependent variable, LX (Table 7) indicated significantly lower probabilities for the independent variables, Time [msec], Time [Ticks], GazeX, GazeY, Diameter, LY, LD, RX, and RY, i.e. the levels of confidence for these variables within this model reached much higher. The F value was relatively higher than typical value.

Table 8. Multiple regression analysis for dependent variable, LY

```
Call:
lm(formula = eye.data$LY ~ `Time[msec]` + `Time[Ticks]` + GazeX +
    GazeY + Diameter + LX + LD + RX + RY, data = eye.data)

Residuals:
    Min      1Q  Median      3Q     Max
-369.70  -23.44   -3.95   14.24  454.06

Coefficients:
                Estimate Std. Error t value Pr(>|t|)
(Intercept)   -1.717e+05  9.207e+03 -18.645  < 2e-16 ***
`Time[msec]`   6.222e-05  2.284e-05   2.724  0.00645 **
`Time[Ticks]`  2.706e-13  1.451e-14  18.646  < 2e-16 ***
GazeX         -2.953e-01  2.304e-02 -12.818  < 2e-16 ***
GazeY          5.516e-01  6.650e-03  82.946  < 2e-16 ***
Diameter      -2.933e+01  3.131e+00  -9.366  < 2e-16 ***
LX             4.852e-02  1.669e-02   2.908  0.00364 **
LD             2.569e+01  2.780e+00   9.244  < 2e-16 ***
RX             2.292e-01  1.004e-02  22.832  < 2e-16 ***
RY             5.354e-01  5.525e-03  96.911  < 2e-16 ***

Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1

Residual standard error: 52.35 on 18797 degrees of freedom
Multiple R-squared:  0.9242,    Adjusted R-squared:  0.9242
F-statistic: 2.547e+04 on 9 and 18797 DF,  p-value: < 2.2e-16
```

The multiple regression analysis for dependent variable, LY (Table 8) suggested significantly lower probabilities of independent variables, Time [msec], Time [Ticks], GazeX, GazeY, Diameter, LX, LD, RX, and RY. Hence, the levels of confidence for these variables were very high.

Table 9. Multiple regression analysis for dependent variable, LD

```
Call:
lm(formula = eye.data$LD ~ `Time[msec]` + Diameter + RX + RD,
    data = eye.data)

Residuals:
      Min        1Q    Median        3Q       Max
-6.83e-11 -1.00e-15  3.00e-15  9.00e-15  2.80e-14

Coefficients:
               Estimate Std. Error    t value Pr(>|t|)
(Intercept)  -4.925e-13  2.395e-14 -2.057e+01   <2e-16 ***
`Time[msec]`  8.639e-20  1.660e-19  5.200e-01    0.603
Diameter      2.000e+00  2.490e-14  8.032e+13   <2e-16 ***
RX            1.782e-17  1.085e-17  1.642e+00    0.101
RD           -1.000e+00  2.448e-14 -4.084e+13   <2e-16 ***

Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1

Residual standard error: 4.982e-13 on 18802 degrees of freedom
Multiple R-squared:      1,     Adjusted R-squared:      1
F-statistic: 4.782e+27 on 4 and 18802 DF,  p-value: < 2.2e-16
```

The multiple regression analysis for dependent variable, LD (Table 9) hinted significantly lower probabilities of independent variables, Diameter, and RD. Hence, the levels of confidence for these variables were very high. Further, the residual standard error was found to be very low but F value was very high.

Table 10. Multiple regression analysis for dependent variable, RX

```
Call:
lm(formula = eye.data$RX ~ `Time[msec]` + `Time[Ticks]` + GazeX +
    GazeY + Diameter + LX + LY + RY, data = eye.data)

Residuals:
    Min      1Q  Median      3Q     Max
-156.22   -7.16   -1.08    4.36  823.45

Coefficients:
                Estimate Std. Error t value Pr(>|t|)
(Intercept)   -4.294e+04  6.246e+03  -6.874 6.44e-12 ***
`Time[msec]`   2.932e-05  1.633e-05   1.796   0.0726 .
`Time[Ticks]`  6.767e-14  9.847e-15   6.872 6.55e-12 ***
GazeX          1.706e+00  1.096e-02 155.631  < 2e-16 ***
GazeY         -1.205e-01  5.494e-03 -21.925  < 2e-16 ***
Diameter       9.239e+00  6.556e-01  14.093  < 2e-16 ***
LX            -6.810e-01  1.087e-02 -62.654  < 2e-16 ***
LY             1.176e-01  5.144e-03  22.857  < 2e-16 ***
RY            -3.046e-02  4.831e-03  -6.306 2.94e-10 ***
---
Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1

Residual standard error: 37.51 on 18798 degrees of freedom
Multiple R-squared:  0.9882,	Adjusted R-squared:  0.9882
F-statistic: 1.971e+05 on 8 and 18798 DF,  p-value: < 2.2e-16
```

The multiple regression analysis for dependent variable, RX (Table 10) implied significantly lower probabilities of independent variables, Time [Ticks], GazeX, GazeY, Diameter, LX, LY, and RY. Hence, the levels of confidence for these variables were very high. Further, the F value was very high.

Table 11. Multiple regression analysis for dependent variable, RY

```
Call:
lm(formula = eye.data$RY ~ `Time[msec]` + `Time[Ticks]` + GazeX +
    GazeY + Diameter + LX + LY + LD + RX, data = eye.data)

Residuals:
    Min      1Q  Median      3Q     Max
-492.97  -16.46    1.19   16.66  739.26

Coefficients:
                Estimate Std. Error t value Pr(>|t|)
(Intercept)    1.024e+05  9.989e+03  10.253  < 2e-16 ***
`Time[msec]`   6.511e-05  2.462e-05   2.644  0.00819 **
`Time[Ticks]` -1.615e-13  1.575e-14 -10.255  < 2e-16 ***
GazeX         -3.466e-01  2.482e-02 -13.969  < 2e-16 ***
GazeY          3.969e-01  7.863e-03  50.474  < 2e-16 ***
Diameter       2.283e+01  3.379e+00   6.756 1.46e-11 ***
LX             4.301e-01  1.772e-02  24.279  < 2e-16 ***
LY             6.222e-01  6.421e-03  96.911  < 2e-16 ***
LD            -2.915e+01  2.996e+00  -9.731  < 2e-16 ***
RX            -6.926e-02  1.096e-02  -6.319 2.70e-10 ***
---
Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1

Residual standard error: 56.44 on 18797 degrees of freedom
Multiple R-squared:  0.9131,	Adjusted R-squared:  0.913
F-statistic: 2.194e+04 on 9 and 18797 DF,  p-value: < 2.2e-16
```

The multiple regression analysis for dependent variable, RY (Table 11) hinted significantly lower probabilities of independent variables, Time [msec], Time [Ticks], GazeX, GazeY, Diameter, LX, LY, LD, and RX. Hence, the levels of confidence for these variables were very high. Further, the F value was very high.

Table 12. Multiple regression analysis for dependent variable, RD

```
Call:
lm(formula = eye.data$RD ~ `Time[msec]` + Diameter + LX + LD,
    data = eye.data)

Residuals:
      Min        1Q    Median        3Q       Max
-6.83e-11 -1.00e-15  3.00e-15  1.00e-14  3.40e-14

Coefficients:
               Estimate Std. Error    t value Pr(>|t|)
(Intercept)  -4.925e-13  2.378e-14 -2.071e+01   <2e-16 ***
`Time[msec]`  8.695e-20  1.661e-19  5.240e-01   0.6005
Diameter      2.000e+00  2.650e-14  7.547e+13   <2e-16 ***
LX            1.816e-17  1.101e-17  1.649e+00   0.0992 .
LD           -1.000e+00  2.448e-14 -4.084e+13   <2e-16 ***
---
Signif. codes:  0 ‘***’ 0.001 ‘**’ 0.01 ‘*’ 0.05 ‘.’ 0.1 ‘ ’ 1

Residual standard error: 4.982e-13 on 18802 degrees of freedom
Multiple R-squared:      1,     Adjusted R-squared:      1
F-statistic: 4.416e+27 on 4 and 18802 DF,  p-value: < 2.2e-16
```

The multiple regression analysis for dependent variable, RD (Table 12) hinted significantly lower probabilities of independent variables, Diameter, and LD. Hence, the levels of confidence for these variables were very high. Further, the residual standard error was found to be very low but F value was very high.

6. Conclusion

By analyzing the eye tracking data for multiple regression modeling, we concluded that there existed significantly a number of suitable and effective models.

At first, the multiple regression model for Diameter variable as dependent on LD and RD variables was considered the best-fit model of multiple regression. The model ensured a very low residual standard error as well. This model elaborated the visual perception during the eye movements, as the focus of resultant fixations depended heavily on both the focus of left eye and the focus of right eye during scene viewing.

Next, the multiple regression models that represented the coordinates, GazeX, GazeY, LX, LY, RX, and RY, were aligned and perceived the cognitive interpretations. As the coordinates of eye fixations that were represented by variables within eye tracking data, had relative dependencies among these left and right eyes along with resultant gaze, hence, the same existed and verified as a valid multiple regression models among these variables.

7. Acknowledgement

At first, our sincere and earnest thanks go to the participants who donated their precious time and efforts during the entire experimentations. We are indebted to their active participations in these research findings. Secondly, we are grateful to the financial supporter; Global IT Consultancy for Innovations without whom the current research work could not be carried out.

8. References

[1] Benjamin D. Yetton, Mohammad Niknazar, Katherine A. Duggan, Elizabeth A. McDevitt, Lauren N. Whitehurst, Negin Sattari, Sara C. Mednick “Automatic detection of rapid eye movements (REMs): A machine learning approach”, Journal of Neuroscience Methods, vol. 259, pp. 72-82, 2016.

[2] R. Zemblys, “Eye-movement event detection meets machine learning”, The 20th International Conference BIOMEDICAL ENGINEERING, Kaunas, Lithuania, pp. 98-101, 2016.

[3] Yingying Jiao, Yong Peng, Bao-Liang Lu, Xiaoping Chen, Shanguang Chen, Chunhui Wang, “Recognizing Slow Eye Movement for Driver Fatigue Detection with Machine Learning

Approach", International Joint Conference on Neural Networks (IJCNN), Beijing, China, pp. 4035-4041, 2014.

[4] Mattias Nilsson, Joakim Nivre, "Learning Where to Look: Modeling Eye Movements in Reading", Proceedings of the Thirteenth Conference on Computational Natural Language Learning (CoNLL), Boulder, Colorado, pp 93–101, 2009.

[5] Michael Burch, Lewis Chuang, Brian Fisher, Albrecht Schmidt, Daniel Weiskopf, Eye Tracking and Visualization: Foundations, Techniques, and Applications, Springer International, 2017.

[6] Aga Bojko, Eye Tracking: The User Experience, a practical guide to research, New York: Rosenfeld Media, 2013.

[7] Andrew T. Duchowski, Eye Tracking Methodology: Theory and Practice, Springer International, 2007.

[8] Jana Holsanova, Discourse, Vision, and Cognition, New York: John Benjamins Publishing Company, 2008.

[9] John M. Henderson and Andrew Hollingworth, "High-level Scene Perception," Annual Review of Psychology, vol. 50, no. 1, pp. 243-271, 1999.

[10] Gufran Ahmad, "Visual Focus of Attention Actively Associates Relevancy in Eye Movements", Journal of Business Theory and Practice, vol. 3, no. 2, pp. 209-223, 2015.

[11] Gufran Ahmad, "Associative Relevance Based Stimulus Shifts Focus in Eye Movements", International Business Research, vol. 8, no. 10, pp. 25-34, 2015.

[12] Gufran Ahmad, "Dynamic Shifts in Visual Focus of Eye Movements Emphasize Associative Relevancy", Research in World Economy, vol. 6, no. 4, pp. 18-28, 2015.

[13] Gufran Ahmad, "Active Changes in Visual Focus of Eye Movements indicate Associative Relevancy Patterns", International Journal of Advancements in Computing Technology, vol. 8, no. 1, pp. 1-10, 2016.

[14] Gufran Ahmad, "Associative Patterns of Eye Fixations Process Decision Making during Scene Perception", International Journal of Engineering and Industries, vol. 8, no. 1, pp. 1-12, 2017.

[15] Lawrence Stark, Stephen R. Ellis, "Scanpaths Revisited: Cognitive Models Direct Active Looking," in Eye Tracking: Cognition and Visual Perception, R. A. M. J. W. S. Dennis F. Fisher, Ed., Hillsdale: Lawrence Erlbaum Associates, 1981, pp. 193-226.

[16] Robert Snowden, Peter Thompson and Tom Troscianko, Basic Vision: An introduction to visual perception, Oxford University Press, 2012.

[17] Simon P. Liversedge, Iain D. Gilchrist and Stefan Everling, The Oxford Handbook of Eye Tracking, Oxford University Press, 2011.

[18] Roger P. G. Van Gompel, Martin H. Fischer, Wayne S. Murray and Robin L. Hill, Eds., Eye Tracking: A Window on Mind and Brain, Elsevier, 2007.

[19] James E. Hoffman, "Visual Attention and Eye Tracking," in Attention, H. Pashler, Ed., London, Psychology Press, 1998, pp. 119-153.

[20] Kenneth Holmqvist, Marcus Nystrom, Richard Andersson, Richard Dewhurst, Eye tracking: A comprehensive guide to methods and measures, Oxford University Press, 2010.

[21] John K. Tsotsos, A Computational Perspective on Visual Attention, London: MIT Press, 2011.

[22] Frans W. Cornelissen, Enno M. Peters, John Palmer, "The Eyelink Toolbox: Eye tracking with MATLAB and the Psychophysics Toolbox", Behavior Research Methods, Instruments, & Computers, Volume 34, Issue 4, pp 613–617, 2002.

[23] Gerhard Krieger, Ingo Rentschler, Gert Hauske, Kerstin Schill, Christoph Zetzsche, "Object and scene analysis by saccadic eye-tracking: an investigation with higher-order statistics", Spatial Vision, Volume 13, Issue 2, pages 201 – 214, 2000.

[24] Hugh Coolican, Research Methods and Statistics in Psychology, Psychology Press, 2017.

[25] Bryan F. J. Manly, Jorge A. Navarro Alberto, Multivariate Statistical Methods: A Primer, CRC Press, 2004.

[26] Fred Pyrczak, Making sense of Statistics: A conceptual overview, Routledge, 2016.

[27] Peter Dalgaard, Introductory Statistics with R, Springer Science+Business Media, 2008.

[28] Michael J. Crawley, The R book, John Wiley & Sons, 2012.

[29] R. Mark Sirkin, Statistics for social sciences, Sage Publications, 2006.

[30] Schuyler W. Huck, Reading Statistics and Research, Pearson Education, 2012.

[31] Joseph F. Healey, The Essentials of Statistics: A tool for social research, Cengage Learning, 2016.

[32] Kevin J. Keen, Graphics for Statistics and Data analysis with R, CRC Press, 2010.

[33] Maria L. Rizzo, Statistical Computing with R, Chapman & Hall / CRC,2008
[34] Jared P. Lander, R for everyone: Advanced analytics and graphics, Addison-Wesley Professional, 2017.
[35] Benjamin S. Baumer, Daniel T. Kaplan, Nicholas J. Horton, Modern Data Science with R, CRC Press, 2017.
[36] Tilman M. Davies, The book of R, No Starch Press, 2016.
[37] Andrea Cirillo, RStudio for R Statistical Computing Cookbook, Packt Publishing, 2016.