**Department of Industrial and Systems Engineering**

**ISE 560, Stochastic Models in Industrial Engineering**

**Course Project Proposal under the Guidance of Professor Dr. Julie Ivy**

# Using Sentiment & Telemetry Data to Predict Customer Satisfaction

Submitted By – Anmol Ratna Pant, Pulkeshian Daruka, Srikant Putrevu, Jimit Shah

## EXECUTIVE SUMMARY

Lenovo is a US$43 billion global Fortune 500 company and a leader in providing innovative consumer, commercial, and data center technology. Presently, it is accelerating its transformation to become a world leader across every part of the business spectrum. Essential to this motive is the need to grasp public expectation and keep its customer base satisfied by providing them with new and exciting products. This has been made possible by gathering their feedback through surveys and acting upon it accordingly. Their latest attempt has been to gauge this public opinion early on in the product life cycle which would give them an edge by making the necessary changes faster and boosting subsequent sales. So, as an alternative to the official NPS, it is looking at the customer reviews available across other online platforms which are available readily.

The methodology of this project involves developing a multiple linear regression model to predict the NPS of given 5 products from the sentiment and NPS data and studying the influence of telemetry data (Battery Life, Driver Health and Wi-fi connectivity) on customer sentiment and NPS.

A Markov decision model was developed to look at the evolution of sentiment over time and to suggest any intervention that would be beneficial. The five components of the MDP (State Space, Action space, Epoch, Transition probabilities and Reward matrix) were defined as per standard assumptions and an optimal policy was derived. As some of the parameters were assumed, they were varied to check the validity of the result. The optimal policy for each state remained the same. Finally, the results are summarized, and further recommendations and improvements are suggested.

## INTRODUCTION

The Lenovo Group is one of the largest technology companies in the world. For a technology company like Lenovo, early and efficient pNPS of its products plays a key role in providing best quality Products and thus, driving customer satisfaction. Net Promoter Score, or NPS, is a measure of customer experience and ultimately predicts business growth. A Net Promoter Score provides companies with a simple and straightforward metric that can be shared with their front-line employees.

Customers are surveyed and asked to rate on a scale of 0-10, the likelihood of recommending the company or brand. Based on their rating, the respondents are grouped as follows:

- **Promoters** (score 9-10): They love the company's products and are loyal enthusiasts who will keep buying and refer others, fueling growth.
- **Passives** (score 7-8): They are satisfied but unenthusiastic customers who are vulnerable to competitive offerings.
- **Detractors** (score 0-6): They are unhappy customer who can damage company's brand and impede growth through negative word-of-mouth.

**Net Promoter Score = % Promoter Customers - % Detractor Customers**

Even though, NPS is the correct measure of customer satisfaction but these surveys are returned very late and hinder the possibility of Lenovo to intervene and repair the possible software issues which if fixed at right time could have increased the NPS score, and therefore the overall sales of the product.

Unlike survey data, sentiment data is available just after the product launch. This sentiment data contains the star ratings and customer sentiment as analysed by the database. The project aims at using the sentiment score available in form of sentiments and stars rating, and thus predict the NPS score of different Lenovo products using this data. NPS score, if predicted correctly, can enable Lenovo to intervene at any time period and resolve the common issues faced by the customers. This can not only increase the final survey score but also increase the sale indirectly. This way individual and overall pNPS can be improved by analyzing just the sentiment data.

## OBJECTIVE OF THE PROJECT

The objective of the Project is to perform the following tasks given below –

- To devise an analytical model using CID sentiment/stars data and predict pNPS Survey scores.
- To find how the Telemetry data influence pNPS Survey Scores and sentiment/star data.
- To predict the pNPS by testing the predictive analytical model on 5 products.
- Lenovo's intervention to improve pNPS for bottom 3 consumer & commercial products.
- To find what impact do the products stars and sentiments have on the analysis.

**METHODOLGY**

The Project is divided into four broad phases –

1. Data Preparation and Analysis
2. Predictive Analytical Modeling
3. Stochastic Optimization using Markov Decision Process.
4. Telemetry Data Analysis

## 1. DATA PREPARATION & ANALYSIS

Lenovo provided us with 3 files containing Web Sentiment Data from Customers across the globe, Telemetry data from Microsoft and actual NPS scores from their customers.

Data Validation and Cleaning was done on the 142223 rows in the Web Sentiment and 48921 rows in the NPS dataset & the data was divided based on product type (Consumer & Commercial Products) to get 142 unique products. For each category, the unique products were filtered out to match with the web sentiment & the NPS Survey data. There were 74 consumer & 68 commercial products identified after initial filtering process. The unique products are shown in Appendix A.

The unique products were defined based on their Star-Sentiment evolution. There were 15 states defined for the analysis with 5 Star ratings (ranging from 1 to 5) and 3 Sentiment evolutions (positive, neutral and negative). The evolution of the star-rating and sentiment for each product was observed on a weekly basis. The figure below explains the evolution of the star-sentiment for the 5 products (these are the same 5 products whose NPS was supposed to be calculated).

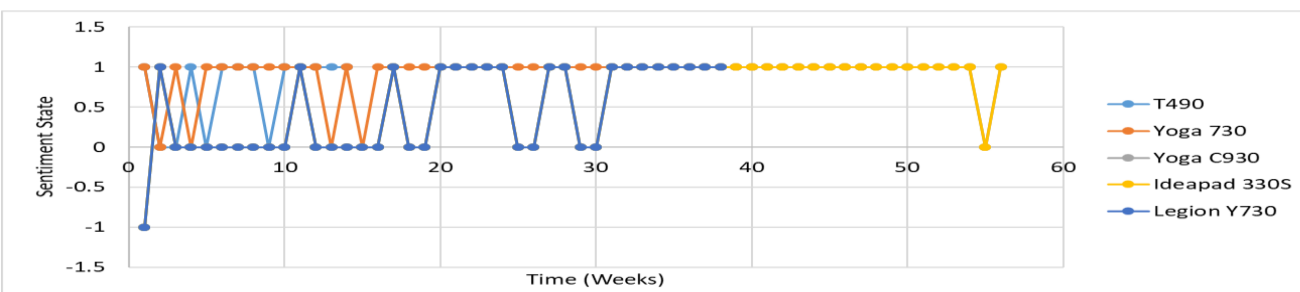



Figure 1: Sentiment Analysis of products based on weekly epochs

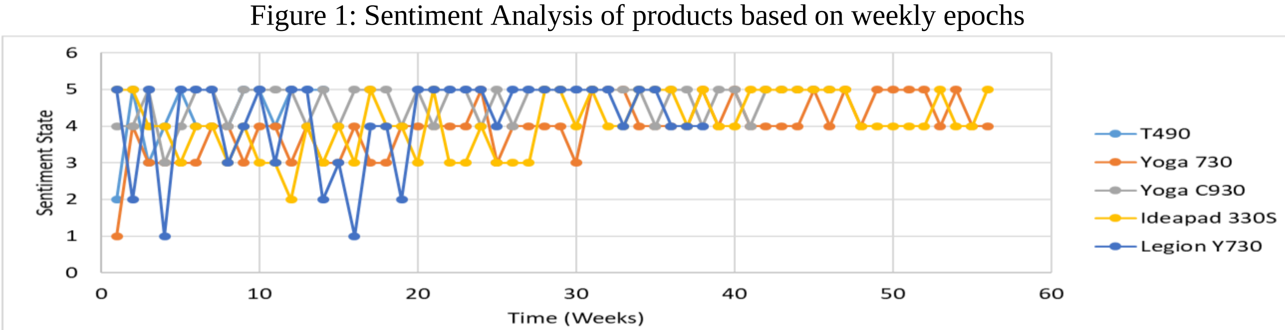



Figure 2: Star Analysis of products based on weekly epochs

## 2. PREDICTIVE ANALYTICAL MODELING

NPS survey and Web Sentiment for each product category was obtained after data filtration & data cleaning and exported to R Studio software for analysis. Analytical Code was developed to match each product in both the datasets (Web Sentiment & NPS Survey). Moreover, calculations of the product sentiment through PSI (Product Sentiment Index) and overall Star Rating for each Product in each category was done and were used as significant factors to formulate the NPS Score. Two Multiple Linear regression models with PSI and Star Rating (Consumer and Commercial Models) were selected among the Liner Regression Model with separate PSI and Star Ratings, Multiple Regression Model with PSI and Star Ratings and Multiple Regression Model with Interactions to predict the pNPS of the required five products. Likewise, the pNPS were predicted for the required five products.

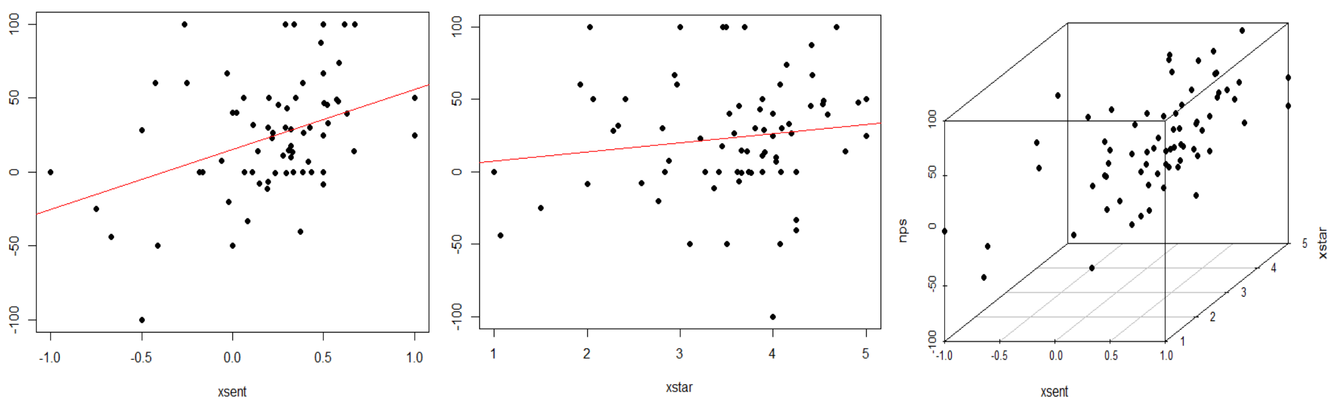



Figure 3: Consumer products pNPS regression fit line

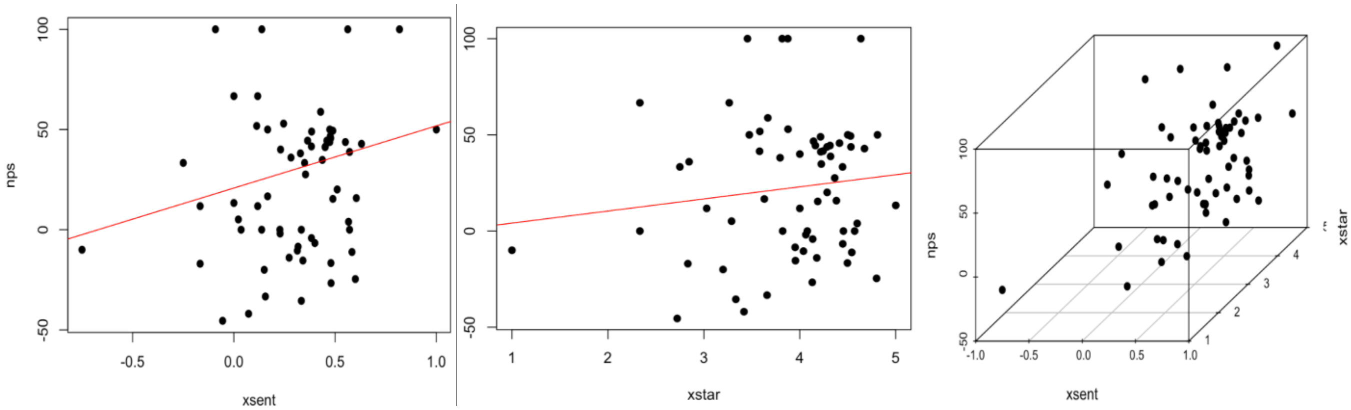



Figure 4: Commercial products pNPS regression fit line

The models gave us the regression fit line with $R^2$ value of 24.25% for consumer products and 4.5% for the commercial products. More information on the model is given in Appendix B. The models were finally used to predict the product NPS values of the required 5 products.

## 3. STOCHASTIC OPTIMIZATION USING MARKOV DECISION PROCESS

The objective of formulating a Markov Decision Model was to understand the evolution of customer sentiment & the appropriate action to improve the NPS Score of the five products. Markov Decision process is a mathematical framework for modeling decisions where outcomes are probabilistic rather than deterministic but can be influenced by the decision maker. This process also suggested best actions to take based on the optimal policy defined.

The components of Markov Decision Process we used for our project are –

- **State Space (S):** The objective was to assess the evolution of sentiment over time and hence the state space comprised of 3 customer sentiments states- **Negative(1), Neutral(2) and Positive(3).**
- **Action Space (A):** This includes the set of possible actions Lenovo could take to impact its NPS score. There are 2 possible actions – **'Do Nothing' and 'Intervene'**
- **Decision Epoch (T):** The epoch refers to the time interval over which decision is made. Since, the data was available for more than a year, we divided our epochs on a **weekly** basis.
- **Transition Probability (P):** This represents the probability of each possible state of the model in the next time period.
- **Rewards (R):** This considers the immediate value of taking an action at each of the given state.

**Assumptions:**

1. The probabilities for the 'Do Nothing' and 'Intervene' actions are kept same for the analysis.
2. The discount factor is assumed to be 0.9 which is used for analysis in Microsoft Excel.
3. Any intervene action by Lenovo invoked the same cost irrespective of state, the system is in.

The Markovian transition probabilities were built by observing the change in sentiments for the given 5 products on weekly epochs. The transition probabilities for each product is shown in Appendix C. This transition probability was assigned random rewards as and when Lenovo intervenes based on the analysis and assumptions from Telemetry Data. The objective function is formulated based on the transition and reward matrix and solved on Excel to predict the improved pNPS. Thus, this Markovian Decision Process helped improve the pNPS of the 5 given products by allocating rewards for each intervention and associating a probability of change in state for these interventions. This helped in analyzing how timely intervention of Lenovo would help improve the NPS score of their products.

## 4. TELEMETRY DATA ANALYSIS

Telemetry Data Analysis is done considering Battery Life and Driver Health Data.

1. **Battery Life:** Data Filtration on both Screen on Battery Life (SOBL) and Hours on Battery Life (HOBL) was done to find the unique products to make the comparisons. Percentage Difference Calculations for 25th ,50th and 75th percentile battery hours from the average of Average HOBL is obtained. This made analyzing the battery performance per model easier. To calculate the effect of battery performance on NPS for each product, calculations for the 25th ,50th, and 75th percentile NPS for each product was carried out as shown in table 1. A Linear Regression Model was built to predict NPS percentile for required products.

Table 1: Battery Health Data

| | | %ile NPS | | | % difference of SOBL data from average HOBL | | |
|---|---|---|---|---|---|---|---|
| **Models** | **Average of Average Hours of Battery Life (HOBL)** | 25th | 50th | 75th | 25th | 50th | 75th |
| 100S | 5.943392799 | 7 | 9 | 10 | -55.9322 | -3.33333 | 5.084746 |
| 720S | 5.603203148 | 7 | 9 | 10 | -32.1429 | -8.92857 | 21.42857 |
| Miix 320 | 6.844745527 | 5 | 8 | 10 | -41.1765 | -4.41176 | 58.82353 |
| ThinkPad X1 Carbon 6th Gen | 7.77405897 | 8 | 9 | 10 | -32.4675 | -6.49351 | 19.48052 |
| Unknown | 7.758732547 | | | | | | |
| X1 Tablet (2nd Gen) | 4.815421135 | 6.75 | 8 | 9.25 | -39.5833 | -18.75 | 12.5 |
| X1 Yoga (3rd Gen) | 6.609767135 | 7.75 | 9 | 10 | -56.0606 | -45.4545 | -31.8182 |
| YB1_X91L | 7.323352846 | 7 | 8 | 9 | -39.726 | -19.1781 | 6.849315 |
| Yoga C930 | 7.321491485 | no nps data | | | -33.0601 | -9.83607 | |

2. **Driver Health:** NPS data filtration into consumer and commercial product categories and data classification of driver class from the driver health data for both commercial and consumer categories is carried out. Driver comments on NPS data for both categories, mentioning issues associated with driver classes is analyzed and their respective pNPS are noted. NPS scores given by customers from the analysis is categorized into Promoters, Neutrals & Detractors. The calculation of the probabilities of Promoters, Neutrals and Detractors from NPS Score is done given that the customer has something to say about a particular driver class. Results are derived from these probabilities for each driver class.

# RESULTS

## 1. Prediction Analysis Results:

The Multiple linear regression model built on the Sentiment and Star-rating gave us the R2 value of 24.25% for consumer products and 4.5% for the commercial products.

```
Call:
lm(formula = nps ~ xsent + xstar)

Residuals:
    Min      1Q  Median      3Q     Max
-87.476 -23.959  -5.159  20.615  82.152

Coefficients:
            Estimate Std. Error t value Pr(>|t|)
(Intercept)   81.008     24.191   3.349  0.00132 **
xsent         79.598     17.696   4.498  2.7e-05 ***
xstar        -20.832      7.533  -2.765  0.00729 **
---
Signif. codes:  0 ‘***’ 0.001 ‘**’ 0.01 ‘*’ 0.05 ‘.’ 0.1 ‘ ’ 1

Residual standard error: 35.7 on 69 degrees of freedom
  (2 observations deleted due to missingness)
Multiple R-squared:  0.2425,	Adjusted R-squared:  0.2205
F-statistic: 11.04 on 2 and 69 DF,  p-value: 6.902e-05
```

```
Residuals:
    Min      1Q  Median      3Q     Max
-59.876 -27.273   4.445  19.649  89.020

Coefficients:
            Estimate Std. Error t value Pr(>|t|)
(Intercept)   20.812     29.786   0.699    0.487
xsent         31.035     22.556   1.376    0.174
xstar         -2.029      8.674  -0.234    0.816

Residual standard error: 34.97 on 61 degrees of freedom
  (3 observations deleted due to missingness)
Multiple R-squared:  0.04569,	Adjusted R-squared:  0.0144
F-statistic:  1.46 on 2 and 61 DF,  p-value: 0.2402
```

Figure 5: Regression model summary for consumer and commercial products

This trained model was used to predict the pNPS and the equation built to calculate this pNPS is mentioned below –

**pNPS $_{\text{consumer}}$ = 81.008 + 79.598 (average of sentiment) + -20.832 (average of star ratings)**

**pNPS $_{\text{commercial}}$ = 20.812 + 31.035 (average of sentiment) -2.029 (average of star ratings)**

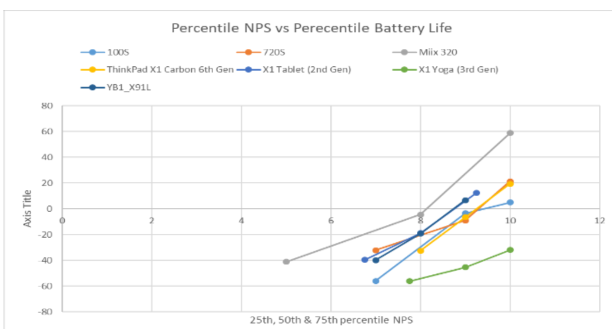
The above formula was used to predict the pNPS of the 5 products which are shown in Table 2.

Table 2: Predicted NPS Score of given 5 products

| PRODUCTS | AVERAGE SENTIMENT | AVERAGE STAR RATING | PREDICTED NPS SCORE |
|---|---|---|---|
| YOGA C930 | 0.5379606 | 4.539674 | 29.30 |
| IDEAPAD 330s | 0.5023374 | 4.302658 | 31.30 |
| LEGION Y730 | 0.3221649 | 3.770619 | 28.11 |
| YOGA 730 | 0.4638525 | 4.225651 | 29.90 |
| T490 | 0.3850267 | 4.144385 | 25.42 |

**2. Telemetry Regression and Analysis Results:**

We plotted the Battery Life data from Table 1 and analysed the percentile NPS against the battery life. The figure 6 shows the results of the analysis. The analysis proved that the NPS and the battery telemetry are closely related with a R-Square value of 49.35% of which regression summary is shown in Appendix D. We estimated the 25$^{th}$, 50$^{th}$, and 75$^{th}$ percentile NPS for all the NPS ratings for YOGA C930 as shown in Figure 7.




| | |
|---|---|
| 25th %ile of all NPS | 8 |
| 50th %ile of all NPS | 8.8 |
| 75th %ile of all NPS | 9 |

Figure 6 : Percentile NPS against Percentile Battery Life

Figure 7 : NPS Prediction for Yoga C930

Driver health data was also analysed to calculate the probability of customer being a promoter, detractor or neutral, given that they have issues associated with the any of the drivers in the driver class. The figure 8 and 9 shows the results of the analysis.

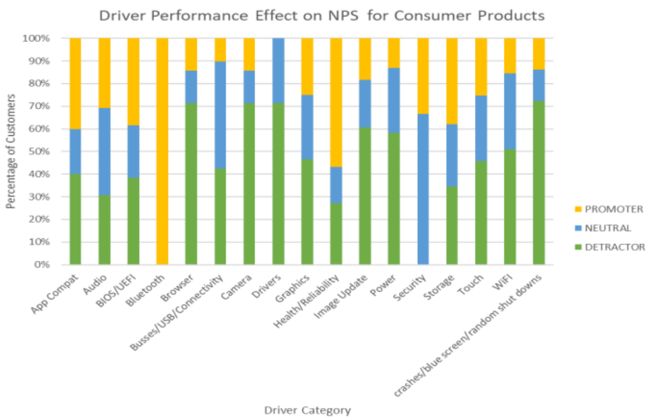



Figure 8: Driver Performance Effect on NPS consumer products

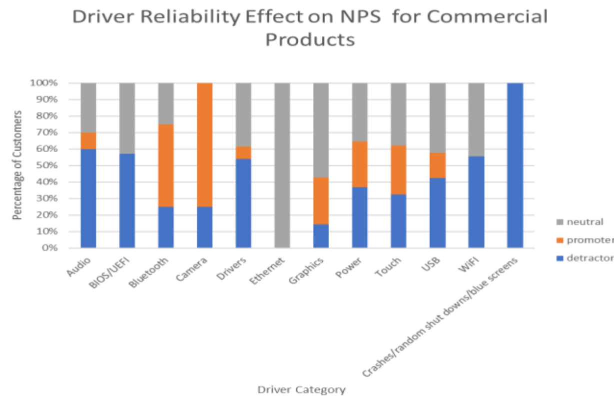



Figure 9: Driver Performance Effect on NPS commercial products

3. Markov Decision Analysis Results:

The Markov Decision process helped in analyzing the pNPS after the intervention is made. The 'Do Nothing' transition probability matrix is computed here based on the evolution of sentiment over time. The 'Intervene' probability matrix has been associated with some rewards to develop an improved NPS policy.

The detailed MDP matrices and calculations are mentioned in Appendix C. The results we got after minimizing the objective function to maximize the product NPS are shown in Table 3 below –

Table 3: Improved NPS of given 5 products devised from Markov Decision Process

| **PRODUCTS** | **YOGA C930** | **IDEAPAD 330s** | **LEGION Y730** | **YOGA 730** | **T490** |
|---|---|---|---|---|---|
| **MDP Predicted NPS Score** | 90.21 | 40.17 | 36.56 | 58.5 | 40.6 |

## RECOMMENDATIONS

After extensively working on the Project, team would like to give the following recommendations:

1. The team suggests that the initial actions for NPS prediction should be made on Multiple Linear Regression Model because of high correlation between sentiment & star rating (96%), thus more weightage to Sentiment should be given. This ensures best prediction for pNPS score through Customer Sentiment and Battery Life data from Telemetry.
2. Secondly, the team recommends on using NPS Score prediction through Markov Decision Process for products through Lenovo's intervention, analyzing Telemetry Data using Linear Programming. Each intervention will help increase NPS and hence taking intervention would be the optimal decision for the company.
3. An exploratory data analysis focusing on each product with unique regression model should be done after collecting proper NPS survey data & sentiment to get accurate results.

4. Lenovo should start a compulsory membership for its customers & should send surveys with a feedback time of one to three months along with offers associated with this time duration to get early feedbacks. It will help to track the customers and take area specific decisions, shorten the time of data collection & thus help to figure out and narrow down problem with each product.
5. The team recommends Lenovo to track the user experience more closely to make decisions on driver intervention because a driver having a huge number of crashes does not necessarily translate to bad user experience. Moreover, third party driver vendors must be communicated accordingly.
6. It must be ensured that the battery performs as close as possible to the battery specifications as it is found directly affects the NPS through battery performance.

   All these Recommendations will help Lenovo improve their product's NPS by solving the problems using optimal policy which will help them improve their products pNPS Score.

**CONCLUSIONS**

Based on the Web Sentiment data and NPS Survey Data for all Geographical regions and for unique products (74 consumer products and 68 commercial products), the team was able to pick out several findings for the objectives of the projects defined and underlying trends from the data. Some of these key findings for the project objectives are as follows:

1. Two Multiple Liner Regression Models – one for 74 unique consumer products and one for 68 unique commercial products were formulated with $R^2$ values of 24.25% and 4.57 % to predict pNPS score for each product.
2. A Multiple Linear Regression Model for percentile NPS on Battery Average HOBL and percentile category is done to predict percentile NPS from which NPS score can be predicted.
3. The pNPS of the five products Yoga C930, Ideapad 330s, Legion Y730, Yoga 730 and T490 are predicted using our Predicted Analytical Model and also predicted using MDP. Also, the predictions were made on NPS Score of Yoga C930 using Battery Life Telemetry data.
4. The bottom three consumer products are Flex 3 14", Yoga 2 Pro and Yoga 3 14" whereas the bottom three commercial products are E560, V110 15" and V130 14". The actions that Lenovo can take to improve NPS are stated under Recommendations.
5. Product Star rating and Sentiment have high correlation values and also it was observed that the Customer Sentiments have high weightage on NPS rather than Star Rating.

**APPENDIX**

**References –**


1. Regression Analysis for predictions - [https://home.ubalt.edu](https://home.ubalt.edu)
2. Wikipedia
3. Introduction to Statistical Learning (Authors – Gareth James, Daniel Witten, Trevor Hastie, Robert Tbshrinani)
4. Applied Statistics & Probability for engineers (Authors – Douglas C. Montgomery, Geroge C. Runger)


## A) Unique Products matching Sentiment & Survey data

| Matching Products | X280 | P300 |
|---|---|---|
| M910 TOWER | N23 CHROME | A285 |
| M710 SFF | M910 TINY | E580 |
| P51 | T470 | V110 15 |
| M83 | X1 TABLET GENERAL | E470 |
| T480 | L470 | E550 |
| P52S | P52 | E590 |
| M73 | THINKPAD 25 | E480 |
| P51S | P1 | V720 14 |
| T580 | P920 | E575 |
| THINKPAD 13 GENERAL | T470S | E570 |
| M93P | T470P | E585 |
| N22 CHROME | L480 | E485 |
| P71 | X270 | V130 15 |
| L580 | T570 | E560 |
| TP YOGA 370 | P72 | V330 15 |
| X380 YOGA | M715 SFF | E460 |
| ONELINK DOCK | TP YOGA 14 | E475 |
| P50 | P50S | E450 |
| M710 TOWER | THINKPAD TABLET 8 | E490 |
| M910Z | TP YOGA 12 | V130 14 |
| T480S | P70 | V330 14 |
| | P40 YOGA | V530S |
| | L390 | |

Figure 1: Unique Commercial Products

| | | |
|---|---|---|
| PRODUCTS | IDEACENTRE Y900 | MIIX 700 |
| CHROMEBOOK C330 | IDEAPAD 100 14 | MIIX2 8 |
| CHROMEBOOK S330 | IDEAPAD 100 15 | MIRAGE SOLO |
| FLEX 2 14 | IDEAPAD 100S 14 | PHAB 2 |
| FLEX 2 15 | IDEAPAD 110 14 | PHAB 2 PLUS |
| FLEX 3 11 | IDEAPAD 110 15 | PHAB 2 PRO |
| FLEX 3 14 | IDEAPAD 110 17 | STAR WARS: JEDI CHALLENG |
| FLEX 3 15 | IDEAPAD 120S 11 | TAB 3 A8 |
| H50-50 | IDEAPAD 120S 14 | TAB 4 8 |
| IDEACENTRE 300 | IDEAPAD 300 15 | YOGA 2 11 |
| IDEACENTRE 310S | IDEAPAD 310 15 | YOGA 2 13 |
| IDEACENTRE 510A | IDEAPAD 320 15 | YOGA 2 GENERAL |
| IDEACENTRE 510S | IDEAPAD 320S 14 | YOGA 2 PRO |
| IDEACENTRE 710 | IDEAPAD 320S 15 | YOGA 3 11 |
| IDEACENTRE 720 | IDEAPAD 500 15 | YOGA 3 14 |
| IDEACENTRE AIO 330-20 | IDEAPAD 510 15 | YOGA 3 PRO |
| IDEACENTRE AIO 510-22 | IDEAPAD 520 15 | YOGA 300 |
| IDEACENTRE AIO 510-23 | IDEAPAD 700 15 | YOGA 500 14 |
| IDEACENTRE AIO 510S | IDEAPAD 700 17 | YOGA 700 11 |
| IDEACENTRE AIO 520-22 | LEGION Y520 | YOGA 710 11 |
| IDEACENTRE AIO 520-24 | LEGION Y520 TOWER | YOGA 710 15 |
| IDEACENTRE AIO 520-27 | LEGION Y740 15 | YOGA 720 13 |
| IDEACENTRE AIO 520S | LENOVO SMART DISPLAY 8 | YOGA 720 15 |
| IDEACENTRE Y700 | MIIX 300 | YOGA 900 GENERAL |
| IDEACENTRE Y710 CUBE | MIIX 630 | YOGA TAB 3 PLUS |

Figure 2: Unique Consumer Products

## B) NPS Prediction Regression Model

```
    Min      1Q  Median      3Q     Max
-58.617 -26.995   3.099  19.637  88.470

Coefficients:
            Estimate Std. Error t value Pr(>|t|)
(Intercept)   14.016      6.556   2.138   0.0365 *
xsent         27.352     16.033   1.706   0.0930 .
---
Signif. codes:  0 ‘***’ 0.001 ‘**’ 0.01 ‘*’ 0.05 ‘.’ 0.1 ‘ ’ 1

Residual standard error: 34.7 on 62 degrees of freedom
  (3 observations deleted due to missingness)
Multiple R-squared:  0.04484,	Adjusted R-squared:  0.02943
F-statistic:  2.91 on 1 and 62 DF,  p-value: 0.09302
```

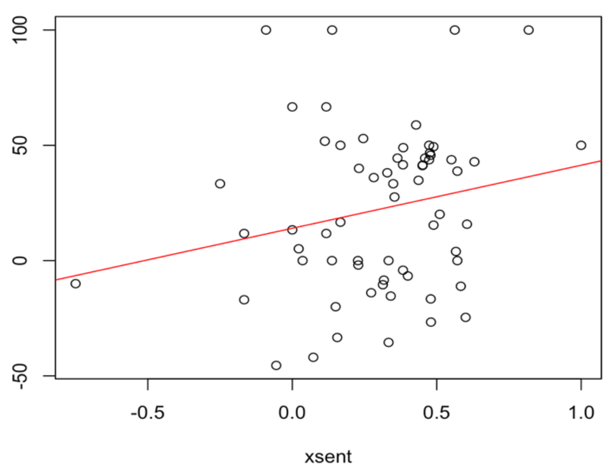



Figure 3: Regression model NPS vs Sentiment Rating for commercial products

```
Residuals:
    Min      1Q  Median      3Q     Max
-61.273 -26.954  -0.857  20.592  80.434

Coefficients:
            Estimate Std. Error t value Pr(>|t|)
(Intercept)   -2.191     24.828  -0.088    0.930
xstar          6.298      6.258   1.006    0.318

Residual standard error: 35.22 on 62 degrees of freedom
  (3 observations deleted due to missingness)
Multiple R-squared:  0.01608,   Adjusted R-squared:  0.0002054
F-statistic: 1.013 on 1 and 62 DF,  p-value: 0.3181
```

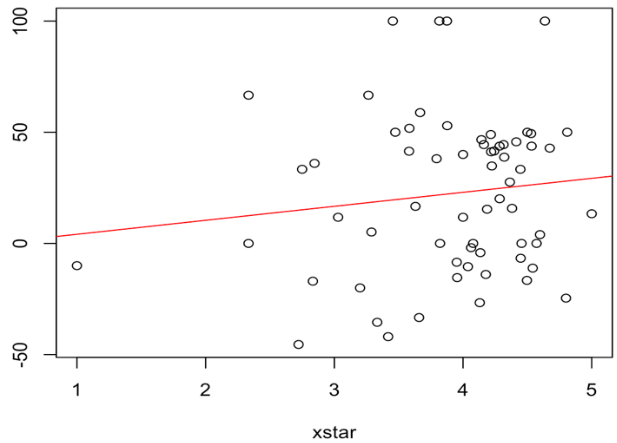



Figure 4: Regression model NPS vs Star Rating for commercial products

```
Call:
lm(formula = nps ~ xsent)

Residuals:
   Min     1Q Median     3Q    Max
-95.09 -28.15  -4.03  24.77  95.30

Coefficients:
            Estimate Std. Error t value Pr(>|t|)
(Intercept)   15.378      4.907   3.134 0.002517 **
xsent         40.573     11.173   3.631 0.000533 ***
---
Signif. codes:  0 ‘***’ 0.001 ‘**’ 0.01 ‘*’ 0.05 ‘.’ 0.1 ‘ ’ 1

Residual standard error: 37.36 on 70 degrees of freedom
  (2 observations deleted due to missingness)
Multiple R-squared:  0.1585,    Adjusted R-squared:  0.1465
F-statistic: 13.19 on 1 and 70 DF,  p-value: 0.0005327
```

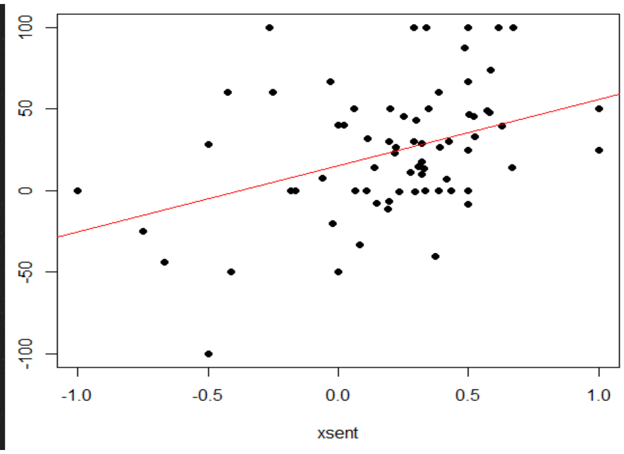



Figure 5: Regression model NPS vs Sentiment for consumer products

```
Call:
lm(formula = nps ~ xstar)

Residuals:
    Min      1Q  Median      3Q     Max
-126.25  -24.06   -1.12   18.02   85.96

Coefficients:
            Estimate Std. Error t value Pr(>|t|)
(Intercept)    1.496     18.646   0.080    0.936
xstar          6.190      5.132   1.206    0.232

Residual standard error: 40.31 on 70 degrees of freedom
  (2 observations deleted due to missingness)
Multiple R-squared:  0.02036,   Adjusted R-squared:  0.006365
F-statistic: 1.455 on 1 and 70 DF,  p-value: 0.2318
```

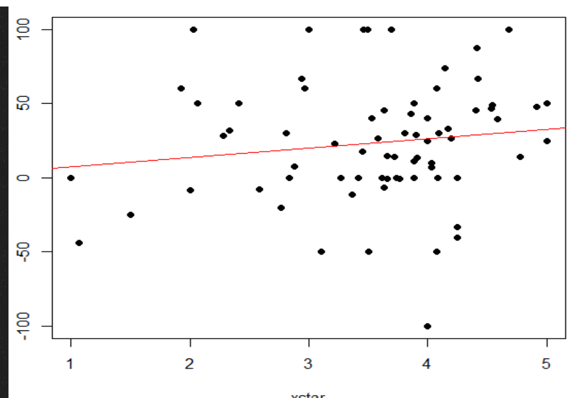



Figure 6: Regression model NPS vs Star Rating for consumer products

## C) Markov Decision Process

### Transition Matrices

| | Negative(-1) | Neutral(0) | Positive(1) |
|---|---|---|---|
| | 1 | 2 | 3 |
| 1 | 0.006173 | 0.018519 | 0.012346 |
| 2 | 0.012346 | 0.166667 | 0.160494 |
| 3 | 0.018519 | 0.15432 | 0.450616 |

Ideapad 330S

| | Negative(-1) | Neutral(0) | Positive(1) |
|---|---|---|---|
| | 1 | 2 | 3 |
| 1 | 0 | 0 | 0.01204813 |
| 2 | 0 | 0.10843372 | 0.12650602 |
| 3 | 0.0060242 | 0.12650601 | 0.62048192 |

Yoga 730

| | Negative(-1) | Neutral(0) | Positive(1) |
|---|---|---|---|
| | 1 | 2 | 3 |
| 1 | 0.013158 | 0.026316 | 0.078947 |
| 2 | 0.013158 | 0.065789 | 0.157895 |
| 3 | 0.092105 | 0.131579 | 0.421053 |

Legion Y730

| | Negative(-1) | Neutral(0) | Positive(1) |
|---|---|---|---|
| | 1 | 2 | 3 |
| 1 | 0 | 0 | 0 |
| 2 | 0 | 0 | 0.06382979 |
| 3 | 0 | 0.08510638 | 0.85106383 |

Yoga C930

| | Negative(-1) | Neutral(0) | Positive(1) |
|---|---|---|---|
| | 1 | 2 | 3 |
| 1 | 0 | 0 | 0 |
| 2 | 0 | 0.071428572 | 0.214285714 |
| 3 | 0 | 0.214285714 | 0.5 |

T490

Figure 7: Transition probabilities based on sentiment for given 5 products

### Reward Matrices

| | Reward Matrix | |
|---|---|---|
| | Do not Intervene(1) | Intervene(2) |
| 1 | -100 | -50 |
| 2 | 5 | 30 |
| 3 | 70 | 100 |

Figure 8: Reward Matrix used to improve NPS for 5 given products

### Objective Function

Max $Z = R_{11}Y_{11} + R_{12}Y_{12} + R_{21}Y_{21} + R_{22}Y_{22} + R_{31}Y_{31} + R_{32}Y_{32}$

Here R represents the rewards associated with each action and state

Subject to –

$(Y_{11} + Y_{12}) - (P_{11}Y_{11} + P'_{11}Y_{12} + P_{21}Y_{21} + P'_{21}Y_{22} + P_{31}Y_{31} + P'_{31}Y_{32})$

$(Y_{21} + Y_{22}) - (P_{12}Y_{11} + P'_{12}Y_{12} + P_{22}Y_{21} + P'_{22}Y_{22} + P_{32}Y_{31} + P'_{32}Y_{32})$

$(Y_{31} + Y_{32}) - (P_{13}Y_{11} + P'_{13}Y_{12} + P_{23}Y_{21} + P'_{23}Y_{22} + P_{33}Y_{31} + P'_{33}Y_{32})$

Here P represents the transition probabilities to each state for ‘Do Nothing’ action and P’ represents the transition probabilities to each state for ‘Intervene’ action.

## D) Telemetry Data

```
Residuals:
    Min      1Q  Median      3Q     Max
-2.4697 -0.6940  0.1604  0.5148  2.0105

Coefficients:
                                   Estimate Std. Error t value Pr(>|t|)
(Intercept)                        8.632840   0.494941  17.442 1.01e-12 ***
BatteryRegression$Battery.Difference 0.033302   0.007985   4.171 0.000574 ***
BatteryRegression$Category         0.069367   0.113798   0.610 0.549772
---
Signif. codes:  0 ‘***’ 0.001 ‘**’ 0.01 ‘*’ 0.05 ‘.’ 0.1 ‘ ’ 1

Residual standard error: 1.014 on 18 degrees of freedom
Multiple R-squared:  0.4935,	Adjusted R-squared:  0.4372
F-statistic: 8.769 on 2 and 18 DF,  p-value: 0.002194
```

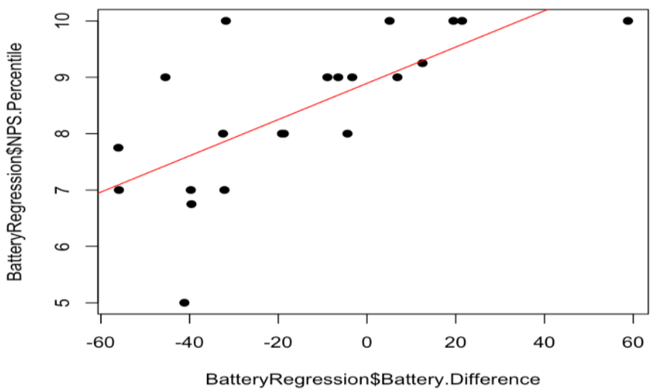



Figure 9: Regression on NPS and Percentile Battery Life