Executive Summary

Sales Representative Net Promoter Score Analysis

"The key to a successful business is not just satisfied customers, but loyal customers" - Shep Hyken. This quote highlights the importance of Net Promoter Score (NPS) as a key performance indicator for businesses. NPS is a management tool used to measure customer loyalty and satisfaction with a brand, product, or service. A high NPS score indicates that customers are more likely to promote a business to others, while a low score highlights areas that require improvement.

Apple Inc. is a prime example of a company that understands the significance of NPS. Apple's NPS score has consistently remained high due to its customer-centric approach and focus on providing a seamless user experience. Apple ensures that its customers not only receive high-quality products but also top-notch customer service. This has helped the company build a loyal customer base that promotes its products through word-of-mouth referrals.

This executive summary aims to address the challenge of identifying the key drivers of a high Net Promoter Score (NPS) among college-educated employees in the software product group. The primary objective is to create a classification model that can effectively predict whether a sales representative will receive an NPS score of 9 or higher. Achieving an NPS score of 9 or above is crucial in indicating high levels of customer satisfaction and loyalty.

Overview

The executive summary is divided into five sections, beginning with the exploratory data analysis. The EDA is divided into two major categories, demographics (section 1) and qualifications (section 2). Demographics includes age, gender and personality, while qualification includes experience, certificates, salary and feedback.

The third section presents the key findings of the analysis along with recommendations, and a surprising discovery. The report then moves on to building a model to predict whether a sales representative will receive an NPS score of 9 or higher. Various models are discussed, and the best model is validated and selected. Finally, an interesting real-world validation is presented to conclude the report

---

1. Analysis on impact of NPS by sales representative demographics

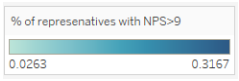
The traits like age, gender, personality, and educational background of a sales rep fall under the category of demographics and can have a significant impact on their performance, which in turn can affect the overall success of the business. Understanding the relationship between sales rep demographics and NPS can help businesses make informed decisions about hiring and training their sales teams to optimize customer satisfaction and loyalty, ultimately leading to increased profitability and growth.

Personality

Personality is an important factor for higher NPS because it influences how effectively they can build and maintain relationships with customers. Certain personalities may be better suited for different types of sales roles, so understanding the personality traits that are most effective for a given role can help organizations select and develop the right sales reps for the job.

For this analysis, the sales representatives are categorized based on their age and gender. Each sales representative will fall in to four categories i.e., young (less than 25), adult (25-35), middle aged (35-45) and senior (45-60). Four personality traits are considered and the percentage of representatives with NPS greater than 9 in each category is reported in the image below.

		Personality			
Gender	**Age category**	Analyst	Diplomat	Explorer	Sentinel
Female	Young	3.17	20.99	22.65	8.25
	Adult	3.55	28.07	28.07	7.45
	Middle Aged	4.11	28.28	31.67	6.10
	Senior	10.32	31.12	28.17	7.88
Male	Young	2.63	22.52	20.49	4.39
	Adult	8.38	26.58	29.13	5.12
	Middle Aged	4.81	28.25	26.58	10.29
	Senior	7.29	28.41	28.26	8.36



('Diplomat' and 'Explorer' are higher in number among the pool of rep's with NPS>9)

Table 1: Percentage of representatives with NPS greater than 9 in each personality category

From table 1 it is quite evident that among the pool of representatives with NPS greater than 9, the sales representatives with 'diplomat' and 'explorer' personality are more in number. This finding suggests that sales representatives with 'diplomat' and 'explorer' personality have a higher chance of receiving an NPS score greater than 9. On the other hand, reps with 'analyst' and 'sentinel' personality types have a lower chance of achieving an NPS score greater than 9.

This could be due to several reasons, including the fact that certain personality traits may align better with the expectations and preferences of customers. For example, customers may prefer sales reps who are more outgoing and personable, which may be characteristics of the 'diplomat' and 'explorer' personality types. Conversely, customers may not respond as positively to sales reps who are more analytical or reserved, which may be characteristics of the 'analyst' and 'sentinel' personality types.

---

2. Analysis on impact of NPS by sales representative qualification

Experience

Experience is a crucial factor in achieving higher NPS, as it can have both positive and negative effects on a sales representative's performance. While some experienced representatives may become complacent and less motivated compared to their younger counterparts, others can leverage their knowledge to gain a deeper understanding of the products or services they are selling. This can lead to improved communication and more effective problem solving, ultimately resulting in higher customer satisfaction. Thus, it is imperative to analyze the experience factor to gain insights into its impact on NPS. The representatives are classified into three categories based on the year of experience. They are entry level (0-2 years), mid-level (2-6 years) and experienced (more than 6 years). The analysis is presented below.

Gender	Age category	Experience: Entry Level	Mid-Level	Experienced
Female	Young	14.32	25.93	
	Adult	18.74	31.76	41.76
	Middle aged	20.85	32.59	36.54
	Senior	20.76	25.63	47.47
Male	Young	16.64	14.88	
	Adult	19.30	27.98	42.02
	Middle aged	20.03	24.15	36.62
	Senior	20.34	25.13	39.47

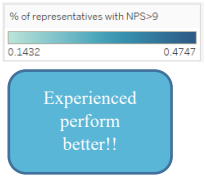


Table 2: Percentage of representatives with NPS greater than 9 in experience category

As shown in Table 2, the percentage of experienced representatives with NPS greater than 9 is higher compared to their entry-level and mid-level counterparts. The deeper blue colour in all rows of the experienced column indicates that experienced representatives perform better irrespective of their age. This can be attributed to the fact that experienced sales reps have likely built a strong network of clients and have honed their sales skills over time, making them more effective in their role. *(Young age range is neglected from the experienced category as minimum of 7 years of experience is required and taking into account that young people cannot reach experienced level at an earlier age.)*

Salary + Feedback:

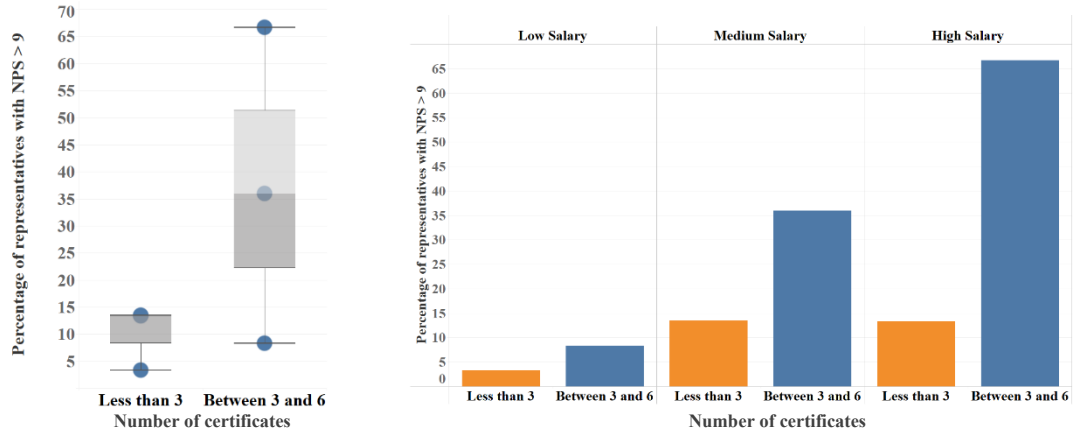
Analysing the sales representative's salary and feedback is important in NPS analysis because they are closely related to job satisfaction and motivation. Feedback can provide insights into areas for improvement, leading to better performance and higher NPS scores. By analysing these factors, companies can identify the key drivers of job satisfaction and motivation, and take steps to improve them, leading to higher NPS scores and overall business success.

Plot 1: Percentage of representatives with NPS >9 against salary and feedback

The above plot 1 gives us several noteworthy observations. It shows that there is a higher percentage of sales reps' with NPS>9 with high salary. There are more percent of high-paid representatives with NPS greater than 9 who outperform their counterparts with lesser salary. Also, when the salary is low the performance of the representative is low. Sales representatives with high salaries have more NPS for several reasons. First, a high salary often reflects a high level of experience and expertise in the industry, which can lead to better performance and higher customer satisfaction. This finding verifies our previous analysis that experienced have more chance of getting NPS greater than 9. Additionally, a high salary can act as a motivator for sales representatives, leading to increased job satisfaction and a higher level of engagement with customers. This finding promotes the company to increase the salary of the sales representatives if they are concerned about increasing the NPS.

Certificates

Certifications can play a crucial role for sales representatives as they can enhance their credibility and perceived level of expertise, thereby increasing customer trust and loyalty. This can ultimately result in higher sales and customer satisfaction. However, it is important to conduct an analysis to understand the impact of the number of certificates held by sales representatives on the NPS score. The summary plots below provide an overview of this relationship



Plot 2: Percentage of representatives with NPS >9 against salary and number of certifications

Plot 2 demonstrates that among the sales representatives who have achieved an NPS score greater than 9, a larger proportion of them possess 3 to 6 certificates. This suggests that having more than three certificates increases the likelihood of a representative achieving an NPS score greater than 9. The third plot reinforces this observation, showing a clear positive trend between the number of certificates held by a representative and their NPS score. Overall, these plots suggest that obtaining certifications is important for sales representatives, as it can lead to higher chance of NPS score greater than 9 and better performance in their role.

---

3. Conclusion:

Based on the analysis of different factors influencing NPS scores among sales representatives, it can be concluded that personality traits, experience, salary, feedback and certifications play a significant role in determining the NPS score. However, it should be noted that age and gender have less impact on achieving NPS greater than 9. The sales representatives with 'diplomat' and 'explorer' personality types tend to have a higher chance of achieving an NPS score greater than 9. Experienced representatives tend to have a higher NPS score than their entry-level and mid-level counterparts. Sales representatives with higher salaries and more certifications tend to have higher NPS scores, which could be due to their level of experience, expertise, and motivation.

Surprising discovery:

One interesting fact identified from the data is that female category and salary are negatively correlated, which means that on average, women tend to earn less than men in the organization. This could be an indication of potential gender-based pay disparities that need to be addressed. It's worth noting that the group of representatives who achieved an NPS greater than 9 had a slightly higher proportion of female representatives.

Recommendations:

Based on the analysis, the following recommendations can be made to improve NPS scores among technical sales representatives:

1. Personality Traits: Identify the personality traits that are best suited for different sales roles and recruit sales representatives accordingly. For this data, "diplomat" and "explorer" personality types are more likely to achieve an NPS score greater than 9, while analyst and sentinel personality types are less likely.

2. Salary: For this data, sales representatives with high salaries have a higher chance of achieving an NPS score greater than 9. A high salary can reflect experience and expertise in the industry, leading to better performance and higher customer satisfaction. Offer competitive salaries to sales representatives to keep them motivated and to improve their job satisfaction and motivation, ultimately resulting in higher NPS scores.

3. Experience: Experience is a key factor in achieving higher NPS scores. Experienced sales representatives have a higher chance of achieving an NPS score greater than 9 compared to their entry-level and mid-level counterparts. Companies should invest in the training and development of their sales representatives to improve their skills and knowledge, ultimately resulting in higher NPS scores.

4. Feedback: Sales representatives with excellent feedback have a higher chance of achieving NPS scores greater than 9. Therefore, it is recommended that companies prioritize representatives with excellent positive feedback if they need NPS greater than 9.

5. Certifications: Obtaining certifications is important for sales representatives, as it can increase their credibility and perceived level of expertise, leading to higher customer trust and loyalty. Sales representatives with 3 to 6 certificates have a higher chance of achieving an NPS score greater than 9.

Implementing these recommendations can lead to higher NPS scores among sales representatives, which can ultimately result in higher customer satisfaction, increased sales, and overall business success.

----------*End of exploratory data analysis*----------

4. MODEL BUILDING

There is a clear need for building a classification model that can effectively predict whether a sales representative will receive an NPS score of 9 or higher. The ability to identify and forecast the performance of reps who are likely to receive high NPS scores can help organizations allocate their resources effectively and achieve their goals efficiently. By leveraging historical data, a classification model can uncover patterns and insights that are not immediately apparent to human analysts, thereby enabling organizations to make data-driven decisions.

3 Base models

The target prediction is to estimate whether the representative will score NPS>9 or not. As the target variable is binary, regression models cannot be used here and so classification models are to be selected. Logistic Regression, Decision Trees, and Naive Bayes are commonly used algorithms for binary classification tasks because of their performance. The models are built based on the appropriateness of the set of predictor variables and various model building features are used to improve the performance. The details of the base logistic regression model are presented below.

Coefficients/Models	Model 1 (Logistic Regression)	Model 2 (Naïve Bayes)	Model 3 (Decision Tree)
Intercept	-7.777e+00***	Does not have coefficients (*Calculates probabilities based on Bayes' theorem*)	Does not have coefficients (*A tree-like model of decisions*)
Age	-1.010e-03		
Years	1.837e-01***		
Personality-Diplomat	1.976e+00***		
Personality-Explorer	1.974e+00***		
Personality-Sentinel	3.860e-02		
Certificates	5.624e-01***		
Feedback	7.026e-01***		
Salary	1.600e-05***		
Female	1.318e-01***		

Note: The asterisk notifies the significance of the particular variable towards the model output

Table 3: Logistic regression model coefficients

It is important to note that in the above table, all the coefficients are significant for the logistic regression model except for age and personality-sentinel. This means that these two variables may not have a significant impact on the NPS score of a customer. As mentioned earlier during exploratory data analysis, it was also observed that the contribution of age towards NPS > 9 was less, and this is reflected in this analysis as well. The other two models have a different approach of predicting the target and do not have coefficients just like the logistic regression model.

Model selection (Accuracy + F1 score):

Model selection based on both accuracy and f1 score is important because accuracy alone may not always provide the complete picture of how well a model performs. Accuracy is a measure of the proportion of correct predictions out of all predictions made by the model. However, when the class distribution is imbalanced, the accuracy metric can be misleading because it can be high even when the model is not performing well for the minority class.

F1 score is a more robust metric that takes into account both precision and recall, and is particularly useful when the classes are imbalanced. F1 score is the harmonic mean of precision and recall, and provides a balanced measure of the model's performance for both classes. For comparing and ranking the three models a combination of accuracy and f1 score is considered. This approach ensures that the selected model not only has a high accuracy but also performs well on both classes, even if they are imbalanced. Therefore, using the combination of accuracy and F1 score provides a more comprehensive evaluation of the models' performance and helps in selecting the best model for the given problem.

K-fold cross validation results:

Results	Model 1(Logistic Regression)	Model 2 (Naïve Bayes)	Model 3 (Decision Tree)
Accuracy (mean)	77.0509	76.19865	50.00
Accuracy (standard deviation)	1.026634	0.6420654	0.023
F1 score (mean)	64.44237	78.55658	52.89751
F1 score (standard deviation)	1.27139	0.5569112	0.3189047
Rank	**2**	**1**	**3**

Table 5: K-fold cross validation results

Based on the k-fold cross-validation table, it is evident that both logistic regression (LR) and Naive Bayes (NB) have achieved a high level of accuracy compared to the decision tree model. However, NB outperforms LR in terms of the f1 score. The ranking is determined based on both accuracy and f1 score, and thus, NB receives the first rank. The standard deviation of both the accuracy and f1 score are minimal. This indicates that the models are stable and consistent in their predictions.

The results obtained from these models are promising, but they may not be sufficient. Therefore, it may be necessary to explore a more powerful model that can better capture the complexity and nonlinearity of the relationship between the predictor variables and the target variable.

Advanced model (Ensemble learning by stacking):

Ensemble learning by stacking is a powerful technique that combines multiple models to improve the accuracy and robustness of predictions. By combining the predictions of multiple models, we can take advantage of the strengths of different models while mitigating their weaknesses. In our case, we have chosen three classification models: logistic regression, decision tree, and Naive Bayes.

Logistic Regression can be used to model the relationship between the rep's characteristics and the target variable, Decision Trees can identify the most important variables for predicting the target variable, and Naive Bayes can be used to classify the reps based on their characteristics. By using a combination of these algorithms, we can build a more robust model that can effectively predict the NPS score of a sales representative.

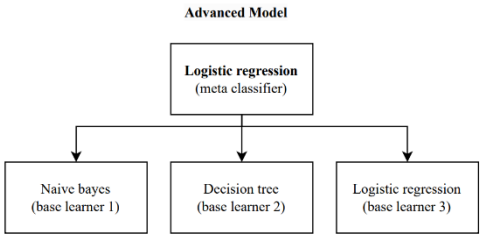


Image 1: Advanced model pipeline

The stacked model takes the outputs of three base models (Naive Bayes + Decision Tree + Logistic Regression) as input and uses a meta-classifier to make the final prediction. The meta-classifier learns to weigh the contributions of each model based on their performance on the validation set. In the advanced model, the meta classifier will be the logistic regression model stacked on top of other three models.

Validation results:

Results	Model 2 (Naïve Bayes) – best base model	Advanced model – ensemble learning
Accuracy (mean)	76.19865	81.66
Accuracy (standard deviation)	0.6420654	0.2343
F1 score (mean)	78.55658	0.8868
F1 score (standard deviation)	0.5569112	0.0454
Rank	2	1

Table 5: Advanced model vs model-2 validation results

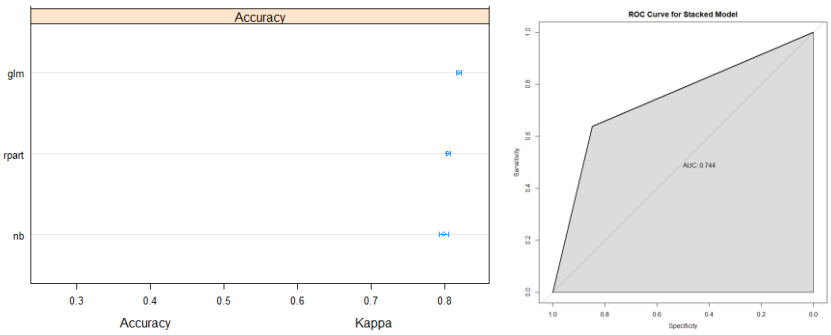


Image 2: Advanced model validation plots

The table 5 compares the performance of the best base model and the advanced model on the test set. The advanced model has significantly outperformed the base model in terms of both accuracy and f1 score. The accuracy of the advanced model is 81.88%, which means that it correctly classified 81.88% of the instances in the test set. This is a significant improvement over the base model's accuracy of 76.19%. The f1 score of the advanced model is 88.68%, which is a measure of the model's accuracy while taking into account both precision and recall. This is a significant improvement over the base model's f1 score of 73.74%. Therefore, it can be concluded that the advanced model is better suited for predicting customer satisfaction compared to the base model.

Referring to image 2, we can observe that all the individual base learners (Naive Bayes + Decision Tree + Logistic Regression) have achieved an exceptional level of accuracy. Furthermore, the stacked model has an AUC of 0.744, which indicates good predictive performance and a good measure of the model's ability to distinguish between positive and negative cases. Therefore, the advanced model is recommended for predicting customer satisfaction based on the given dataset.

Model features:

Data cleaning & sub-setting | Up-sampling minority class | k-fold cross validation | Accuracy +F1 score validation | Ensemble learning | AUC

End of model building

5. Interesting real-world validation

At the end of the exploratory data analysis, we concluded that an experienced female representative with a "diplomat" or "explorer" personality and a high salary would have a higher probability of getting an NPS greater than 9. So, let's verify this conclusion using the advanced model by two real-world scenarios. We will use the advanced model to predict the NPS scores for both representatives and compare them to see if they align with the conclusions drawn from our exploratory data analysis.

Scenario 1: A 35-year-old female representative with 10 years of experience in software, a college degree, 6 certificates, and a feedback score of 4. Her salary is $150,000, and her personality type is Diplomat.

Scenario 2: A 30-year-old male representative with 1 year of experience in software, a college degree, 1 certificate, and a feedback score of 1. His salary is $70,000, and his personality type is Analyst.

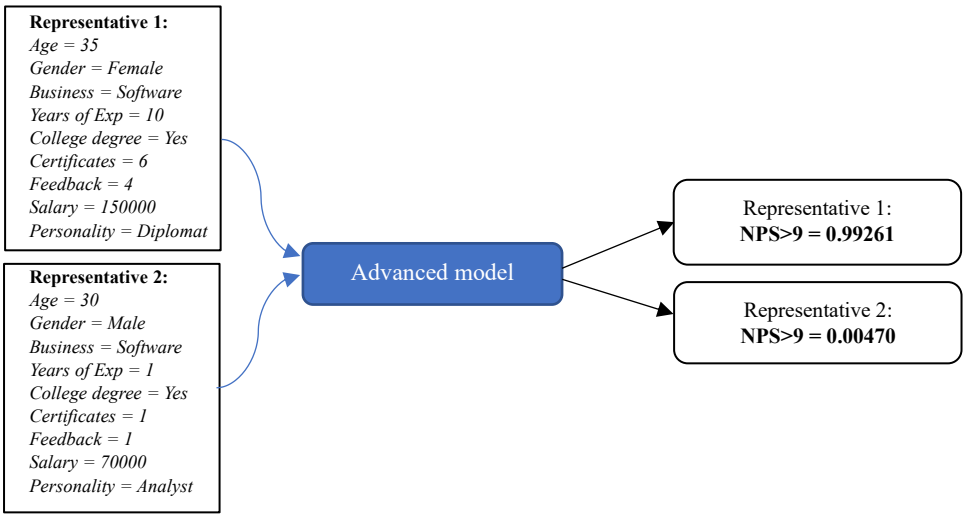


Image 2: Interesting real-world validation

The results from the predictions using the built model indicate that the representative-1 has a high chance (99%) of achieving a NPS greater than 9, while the representative 2 has a very low chance (0.4%) of achieving the same. It should be noted that the two cases chosen were extreme in terms of the variables, which is why there was such a large difference in the percentage values. However, this does support our initial conclusion from the exploratory data analysis that an experienced female representative with a diplomat or explorer personality and a high salary has a higher chance of achieving a NPS greater than 9. It also suggests that variables such as personality, feedback, certificates, and salary may have a significant impact on the NPS, as they were important predictors in the advanced model.

----------*End of report*----------

NPS SCORE

Project GitHub: *https://github.com/johnmelwin/SalesNPSAnalysis*
Tools used: *Excel, R, Tableau*

RIT

Analysed by:
Name: *John Melwin Richard*
Email: *jj5603@rit.edu*