Banking Variable Rate Annuity Product: Model Assessment and Prediction

Alex Rodgers, Brett Laderman, Cameron Houston, Gina Barbagallo, Jason Wang

September 20, 2024

COMMERCIAL BANKING, CORP

Table of Contents

Banking Variable Rate Annuity Product: Model Assessment and Prediction

Overview

The Bank, acting through the department of *Customer Services and New Products,* requested predictions on customers who will purchase a variable rate annuity product.

We used our previous model of 14 predictor variables with one interaction. Our model's concordance was 79.98%, correctly assigning a higher probability to customers who purchased the insurance product compared to those who did not 79.98% of the time. We identified the coefficient of discrimination, optimal cutoff, and prediction accuracy to better gauge performance for our business case.

In this proposal we cover data used, model characteristics, and overall performance. Using this information, we provide detailed recommendations to the Bank about how and when to use this model.

Methodology and Analysis

In this section we discuss the concordance percentage, coefficient of discrimination, ROC curve, and K-S statistic of our selected model. Table 1 displays the model variables and their corresponding p values.

Table 1: Forward Selection P Values

VARIABLE	P VALUE
Savings account balance (SAVBAL_BIN)	8.01E-129
Checking account balance (DDABAL_BIN)	5.63E-60
Certificate of deposit account balance (CDBAL_BIN)	2.58E-39
Money market account balance (MMBAL_BIN)	2.37E-23
Number of checks written (CHECKS_BIN)	5.76E-20
Total ATM withdrawal amount (ATMAMT_BIN)	9.09E-10
Number of teller visit interactions (TELLER_BIN)	1.93E-08
Indicator for credit card (CC)	1.53E-07
Indicator for checking account (DDA)	1.12E-05
Installment loan account balance (ILSBAL_BIN)	5.64E-05
Indicator for investment account (INV)	1.15E-04
Interaction (IRA: DDA)	3.13E-04
Indicator for mortgage (MTG)	6.62E-04
Number of insufficient fund issues (NSF)	9.57E-04
Indicator for retirement account (IRA)	8.80E-01

Table 1 lists variables and interactions significant at our cutoff value of 0.002, including the main terms that were not significant. Each variable is used in our model for client product purchases.

Concordance Percentage

The first probability metric we looked at, concordance percentage, benchmarked at 79.98%. This means in nearly 80% of cases it correctly assigned a higher probability to customers who purchased the insurance product compared to those who did not.

Discrimination Slope

The discrimination slope helps assess how well the model separates insurance purchasers from non-insurance purchasers. Figure 1 shows the predicted probability bins and the distribution of individuals who did and did not purchase the insurance.

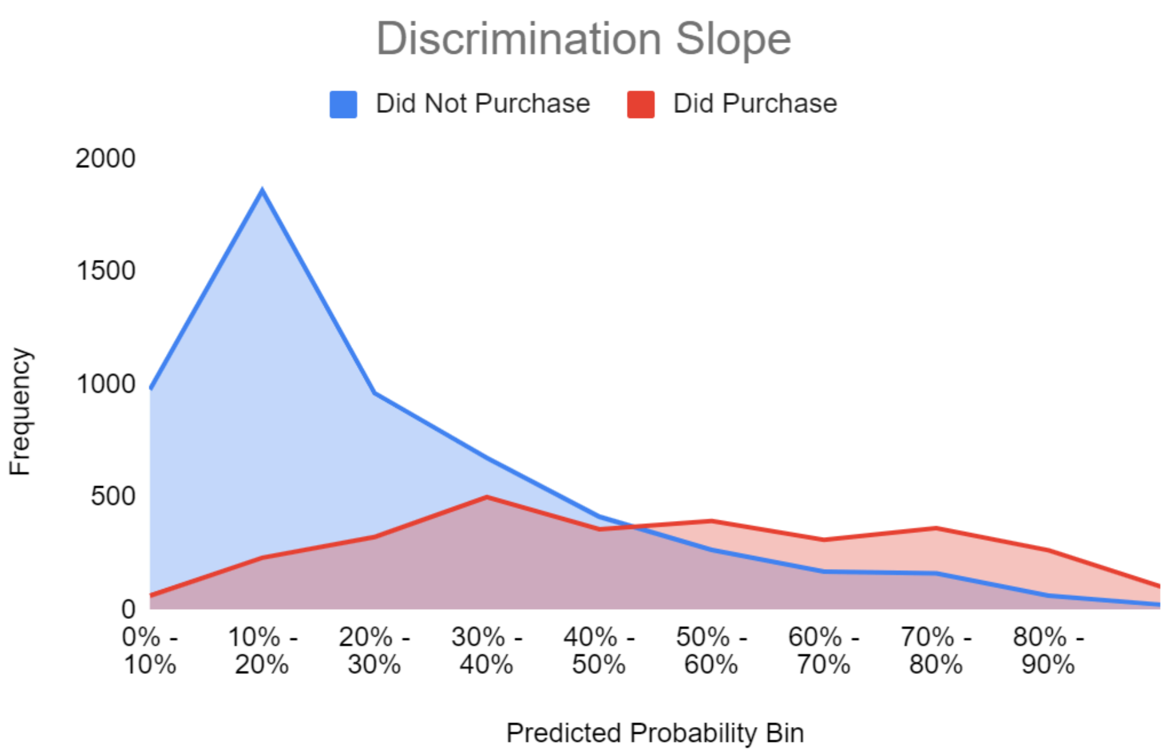

Figure 1: Discrimination Slope

The model's coefficient of discrimination was 0.246, meaning the average predicted probability for purchasers is 0.246 higher than the average predicted probability for non-purchasers.

ROC Curve and K-S Statistic

The curve in Figure 2 shows our model on the data, with Youden's Index showing a cutoff of 0.297 maximizing the combination of the true positive rate and false positive rate. This cutoff yields a K-S statistic of 0.417. This means that if our model is used to predict whether or not a customer will purchase the product, any time it gives a probability of .297 or higher that should be treated as a prediction that the customer will purchase the product, and any time it gives a probability lower than .297 that should be treated as a prediction that the customer will not purchase it. This is the cutoff we recommend using, provided that the cost of false positives is equal to the cost of false negatives.

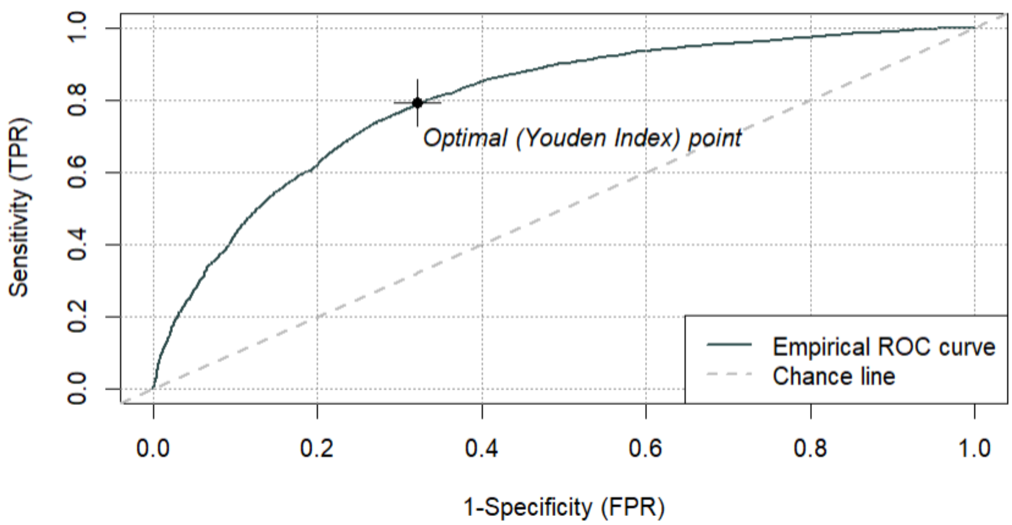

Figure 2: ROC Curve

The curve in Figure 2 shows our model on the data, with the optimal point being at the best true positive rate with the least amount of false positives.

Results and Recommendations

We wanted to see how well the model predicted product purchases by clients not included in the training data. We found that our best accuracy is 73%. The predicted probability that this occurs at (the optimal cut-off) is defined as 0.42.

We used a confusion matrix to look at the exact amount of accurately predicted values compared to what they were initially classified as, as seen in Figure 3**.**

		Actual Values	
		Did not purchase	Made a purchase
Predicted Values	Did not purchase	1163	362
	Made a purchase	219	380

Figure 3: Confusion Matrix

The model accurately predicts 84% of the clients who would not purchase the product and 51% of the clients who would purchase the product, shown in Figure 3 .

Our analysis revealed that our model could detect a client who purchased the product nearly twice as effectively as random selection shown in Figure 4. While evaluating other attributes of the clients could allow us to increase our true purchase predictions, we are encouraged by the precision of 84% when predicting clients that would not purchase the product. This level of precision is insightful when considering marketing techniques and overall underscores our model.

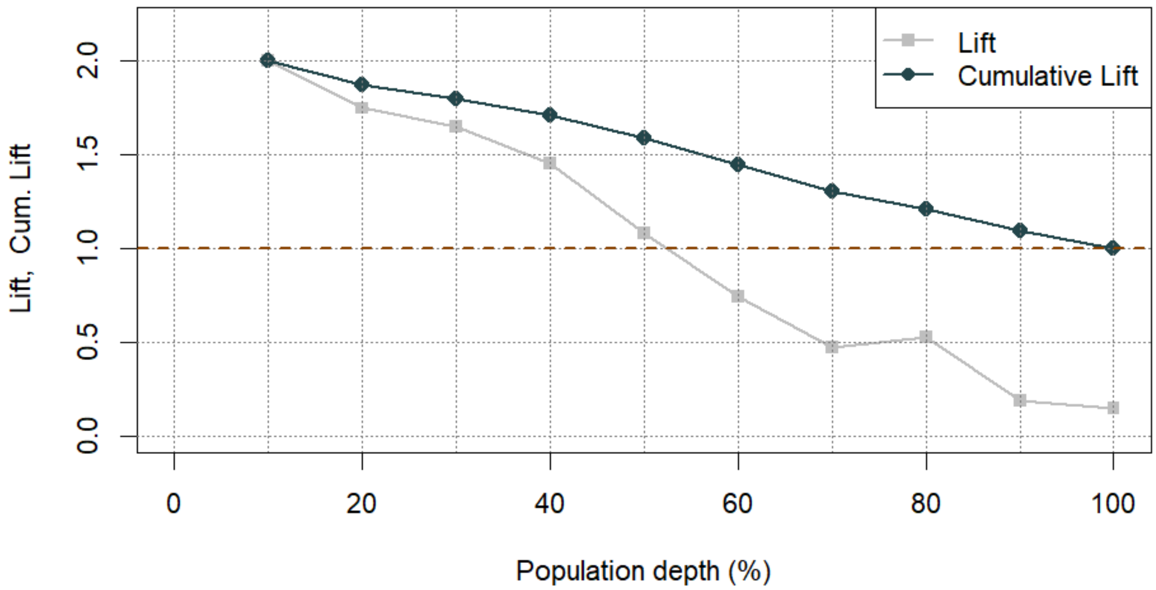

Figure 4: Lift Chart

Figure 4 is our lift chart that shows the percentage of customers who purchased the product from our model compared to a random selection of customers.

Since our model was built around understanding account balances, we recommend that rather than encouraging all clients to open more accounts, the Bank focuses its marketing efforts on clients with larger account balances, specifically in secondary accounts. Targeting those with higher balances will increase the number of clients who purchase the product, optimizing your marketing strategy.

Conclusion

The Bank tasked us with predicting which customers will purchase variable rate annuity products. We investigated, reported, and interpreted various accuracy measures from our model. Our model correctly assigned a higher probability to customers who purchased the insurance product compared to those who did not 79.98% of the time.

Our model identified purchasers almost twice as effectively as random chance, with 84% precision—meaning 84% of predicted buyers actually made a purchase. While adding more attributes could improve accuracy, this strong precision already supports more targeted marketing strategies.

We recommend that the Bank focus its efforts on clients with larger balances in secondary accounts. Targeting those with higher balances is likely to increase the number of clients who purchase the product.

Appendix

Table 2: Gains Table

Bucket	Obs	CObs	Depth	Resp	CResp	RespRate	CRespRate	CCapRate	Lift
1	212	212	0.1	148	148	0.698	0.698	0.199	1.998
2	213	425	0.2	130	278	0.61	0.654	0.375	1.747
3	212	637	0.3	122	400	0.575	0.628	0.539	1.647
4	213	850	0.4	108	508	0.507	0.598	0.685	1.451
5	212	1062	0.5	80	588	0.377	0.554	0.792	1.08
6	212	1274	0.6	55	643	0.259	0.505	0.867	0.743
7	213	1487	0.7	35	678	0.164	0.456	0.914	0.47
8	212	1699	0.8	39	717	0.184	0.422	0.966	0.527
9	213	1912	0.9	14	731	0.066	0.382	0.985	0.188
10	212	2124	1	11	742	0.052	0.349	1	0.149