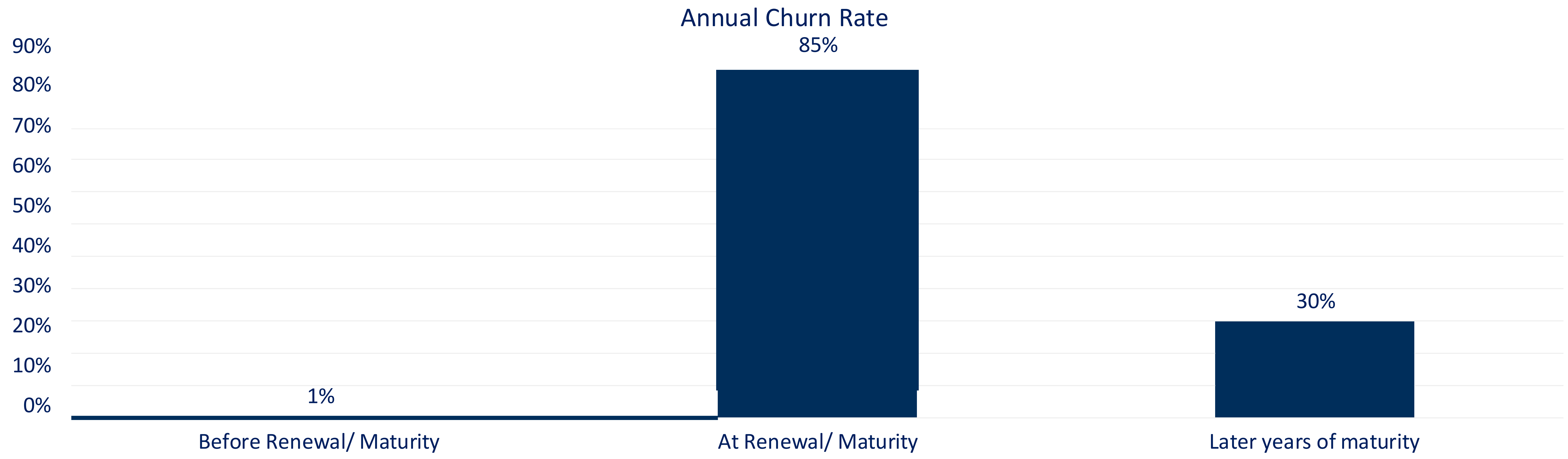
Churn Modeling – Annuity Insurance Industry

Short Summary on Annuity/ Accumulation Products & Churn Behavior

- Annuity/ Accumulation are financial products that offer guaranteed return while you save for retirement
- It comes with different year options like 3 Year Annuity, 5 Year etc. which just means you place money in the account and it grows with guaranteed interest rate for the years you bought the policy for and then the policy renews/ matures with either lower guarantee rate or the same, depending upon company's strategies
- The policy has a surrender penalty if you want to surrender before the maturity/ renewal which usually triggers extreme churn event at the maturity/ renewal year as shown in the chart below
- This product is very similar to a CD at banks but usually has higher rate offered. There is another product called Fixed Indexed Annuity with offers guarantee rate with some indexed credits based on the index performance (like s&p , blackrock fund etc). It also offers downside protection where your money will not go negative unlike when you place your money in a mutual fund



Problem Statement

- XYZ Annuity company has a rich data set of surrender/ churn activity on the annuity products which are being predicted using judgement-based approaches & not using data
- The models are plethora of complex tables of churn rates which are performing poorly across products & hard to maintain
- There are variety of factors which were not included in the current model in place
- The data is a collection of multiple disparate data sets and block of businesses, each with different structure

Approach

- We gathered (from multiple databases) , normalized, joined and validated the data sets to facilitate answering basic question about surrender behavior
- In course of this project, we examined the roles of various factors like renewal years for the contract, contract structures, distribution firms and other demographic / policyholder features in explaining the levels and variations in churn rates
- We used a tree-based bagging based approach (Random Forest) and Feed Forward Neural Network to learn complex features predicting surrenders

Results

- We replaced the legacy judgement based model which had many complex tables of surrender rates per product with a unified model fitted to the data.
- We were able to extrapolate surrender on areas in retail products where we did not have much data on, using data from older products which encouraged data driven decisions
- We developed a model to quantify the influences of various factors on surrender in order to predict it with a material greater degree of accuracy and explanatory power
- We proposed some strategy to improve or drive various Pricing, Finance, Sales, Product Development & Investment strategies

Features in the Proposed Model

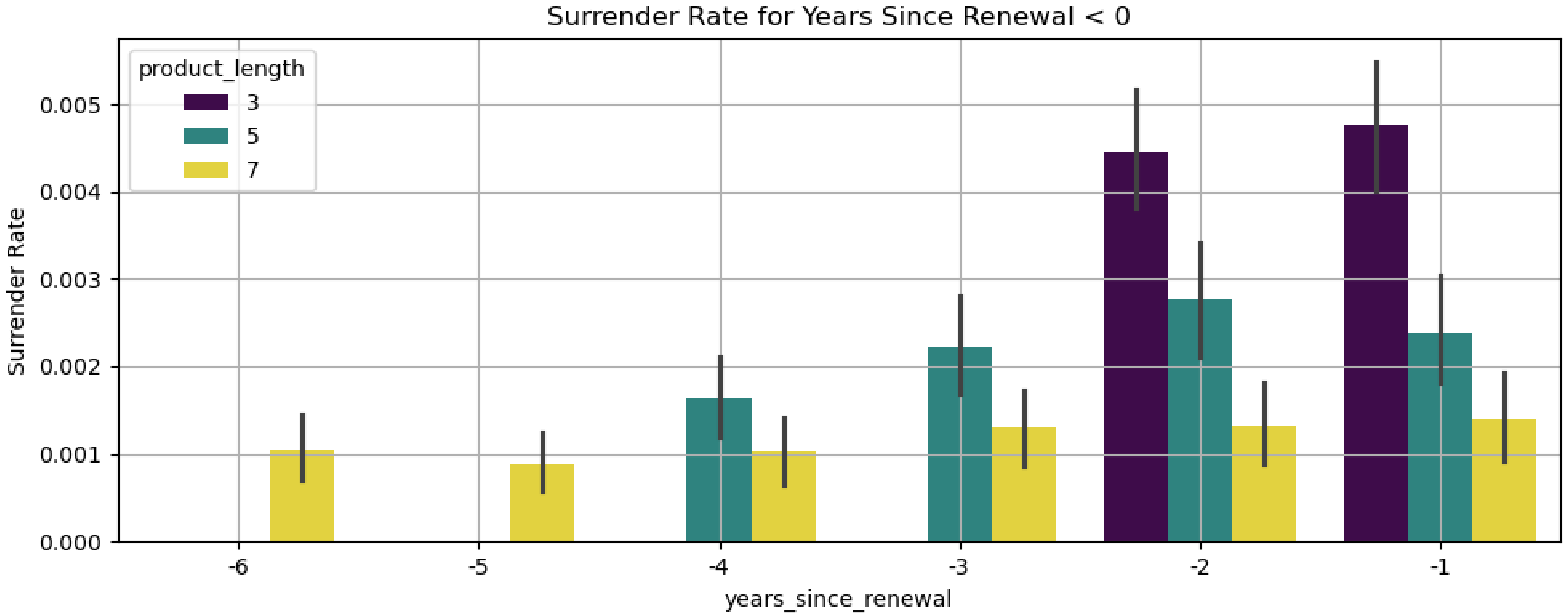
- Here are some key features which will be used in the churn model. These features were designed after surveying product teams.
- Number of parameters will be reduced from ~600 surrender rates to 1 model !

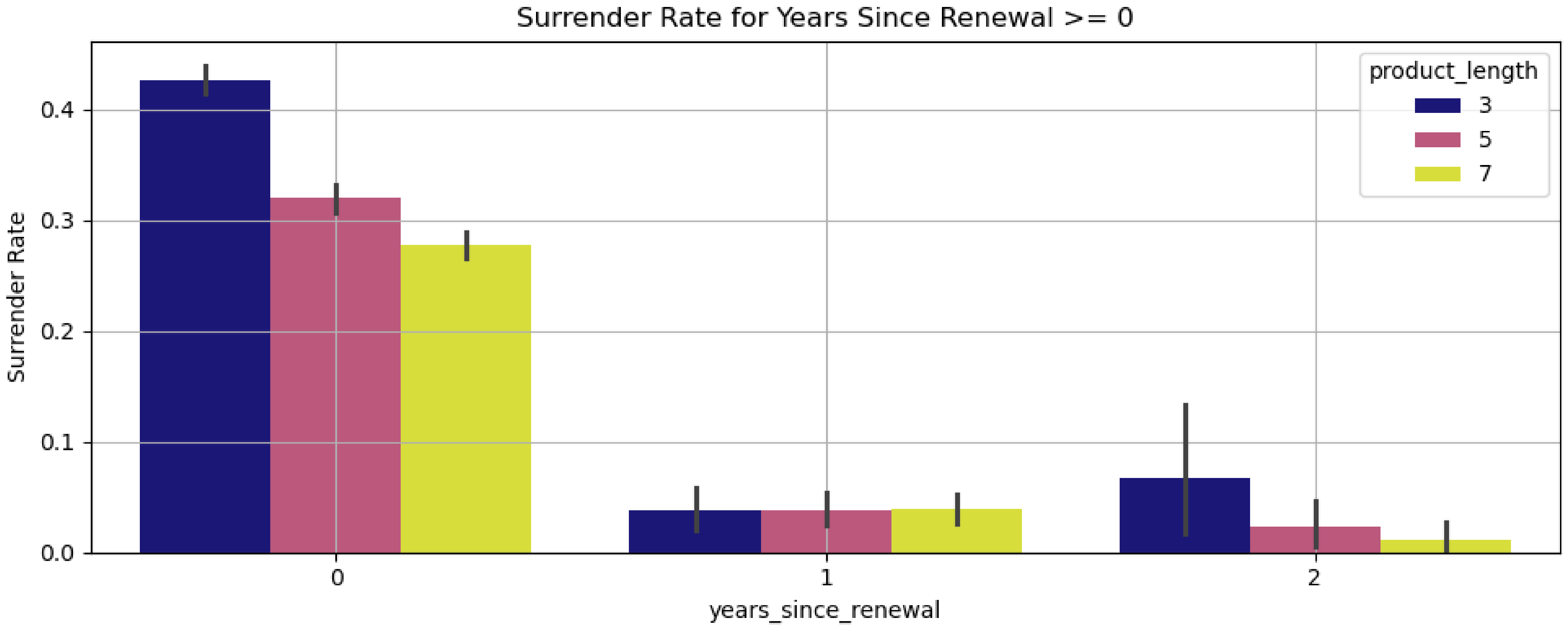
Variable	Definition/Example	Impact to Churn (Intuition)	Variable Type	Category
Years since renewal/maturity	For 3-Year Annuity, 4th year = 1st renewal year	Churn peaks in 1st renewal year, then declines.	Categorical	Calculated
Annuity vs Indexed Annuity	Indexed products tied to S&P, offering fixed + indexed growth with downside protection.	Indexed products have lower surrenders.	Categorical	Known at Issue
Length of product	Product duration (e.g., 3-Year Annuity = 3).	Longer products churn less due to high surrender charge and commitment	Categorical	Known at Issue
Line of business	Reinsured or Retail products.	Churn varies by product type.	Categorical	Known at Issue
Crediting Rate	Rate offered to policy (e.g., 1.5%, 3%).	Higher rates reduce churn.	Numeric	Dynamic
5-Year Treasury Rate	Government 5-year bond yield.	Higher rates increase churn.	Numeric	Dynamic
Attained Age	Current age of policyholder.	Older age increases churn in renewal year; decreases elsewhere.	Numeric	Calculated from DOB
Tax Status	Qualified (pre-tax) or Non-Qualified.	Non-Qualified policies have higher surrenders.	Categorical	Known at Issue
Policy Month	Month 1-12	Higher surrenders in early months.	Categorical	Calculated
Distributor Group	Sales channel (e.g., Bank, Broker-dealers).	Churn varies by distributor.	Categorical	Known at Issue
Account Value Size	Current contract value (e.g., 0-100K, 100K+).	Larger values increase churn in renewal year; decrease elsewhere.	Numeric	Calculated
Our own new money rate	Policies churn to buy new products with us.	Higher rates increase churn.	Numeric	Dynamic

Exploratory Data Analysis

Surrender Rate by Key Variable - Renewal Years

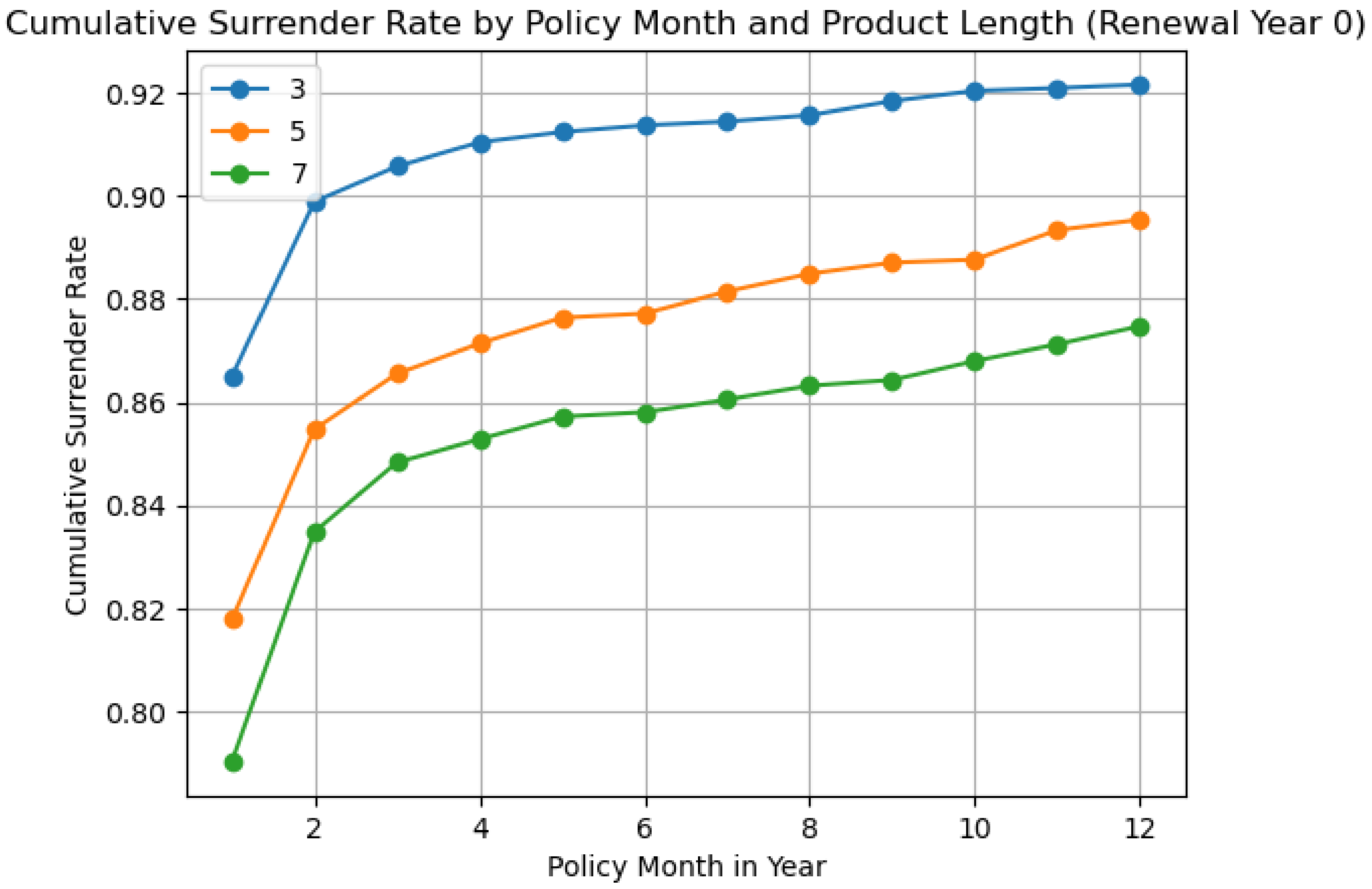
Surrender Rate is expected to be highest in renewal year and lowest before renewal year due to surrender charge penalty.
Surrender Rate will reduce with length of product due to higher surrender charge and financial commitment on longer products.





Cumulative/ Hazard Surrender Rate – Renewal Year 0

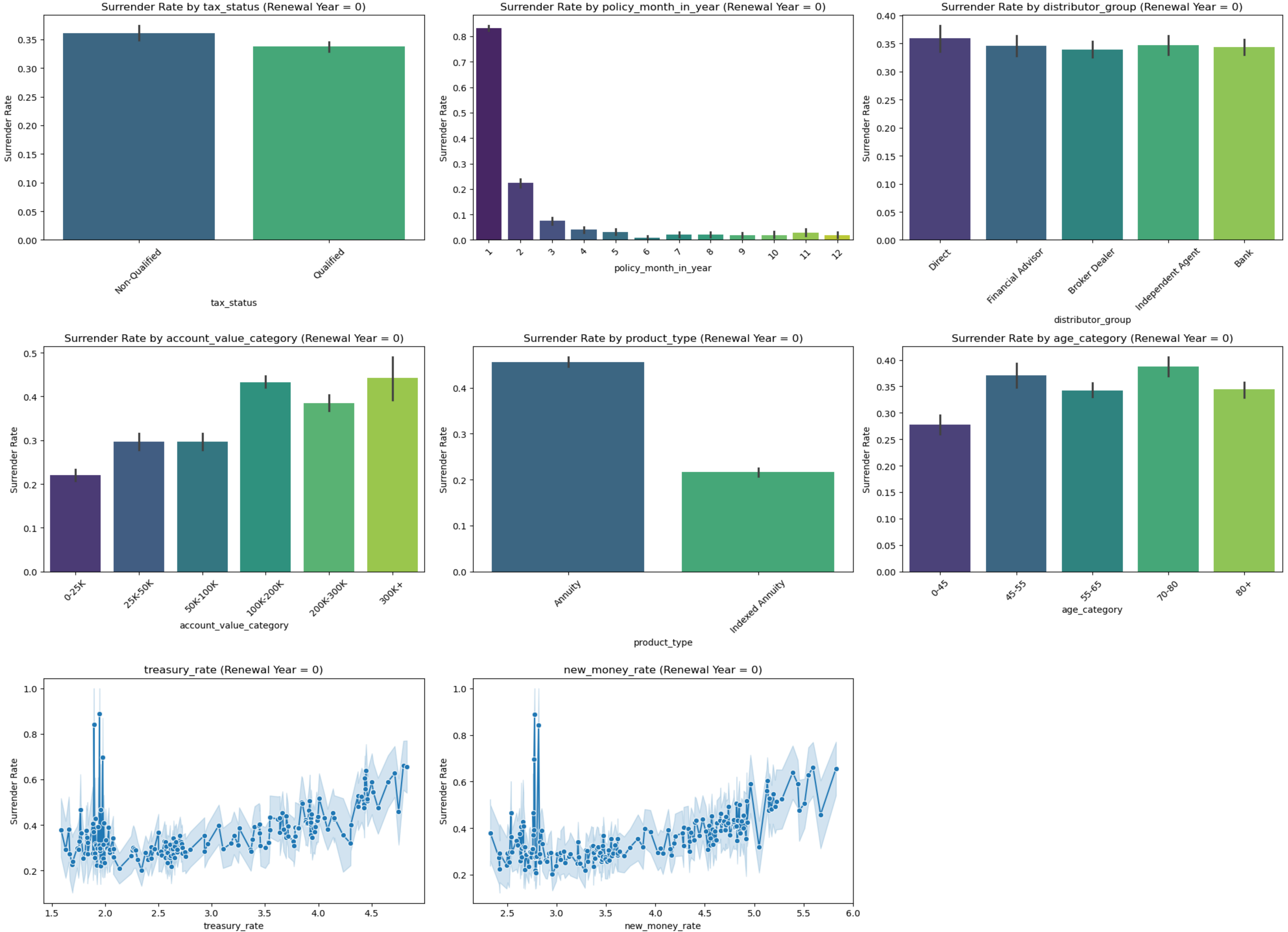
Hazard surrender rate during renewal year (most impactful year to the churn) suggest increasing surrender with decreasing product length.



This chart is useful for pricing teams to run back of envelope Internal Rate of Return hit/ improvement to business given this data, before incorporating the final model we propose into their systems.

Surrender Rate by Other variables – Focusing Renewal Yr 0

Aggregated surrender rate by other variable reveal interesting correlation phenomenon, surrender rate is highest in policy month 1.
Its important to note that surrenders are complex phenomenon, and we should not conclude from aggregated surrender rate as there are lot of confounding/ interaction variables
Hypothetical e.g. surrender rate for Annuity product is higher than Indexed probably, which could be due to fact that we are selling the annuity product to age group 70-80 and they churn more. Hence we would like to learn these complex behavior using a machine learning model.



Model Training & Evaluation

Methodology

1. Data Preparation

- **Dataset**: Historical policy data (since 2007) with features like policy_date, product_type, crediting_rate, account_value, renewal year etc.
- **Preprocessing**:
  - Converted policy_date to datetime and sorted data chronologically.
  - Split data into **training (80%)** and back **testing (20%)** sets based on time.
  - Missing values are minimal in this data set as its clean data set
  - Encoded categorical variables (e.g., product_type, tax_status) using LabelEncoder.
  - Normalized numerical features using StandardScaler.

2 . Model Training

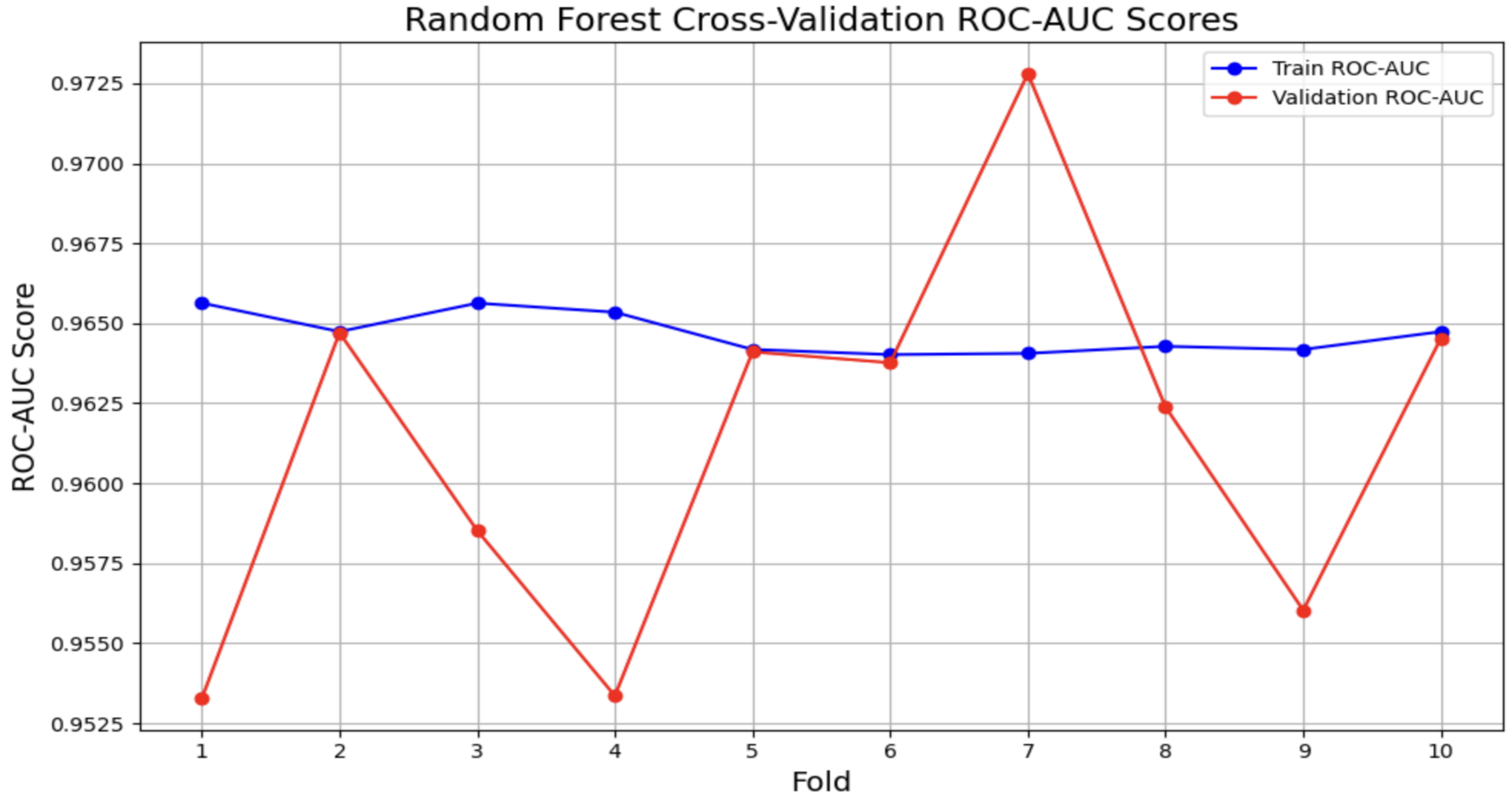
- **Random Forest Classifier**:
  - Trained with 100 trees, max depth of 5, and 10-fold cross-validation.
  - Evaluated using **ROC-AUC** scores for both training and validation sets.
  - Cross-validation results should show consistent performance across folds.
- **Neural Network (FeedForwardNN)**:
  - Architecture: 2 hidden layers (64 and 32 neurons) with **ReLU activation**, **Batch Normalization**, and **Dropout** for regularization.
  - Trained for 20 epochs using **Binary Cross-Entropy Loss** and **Adam optimizer**.
  - Monitored training and validation loss, along with **ROC-AUC scores** for each epoch.

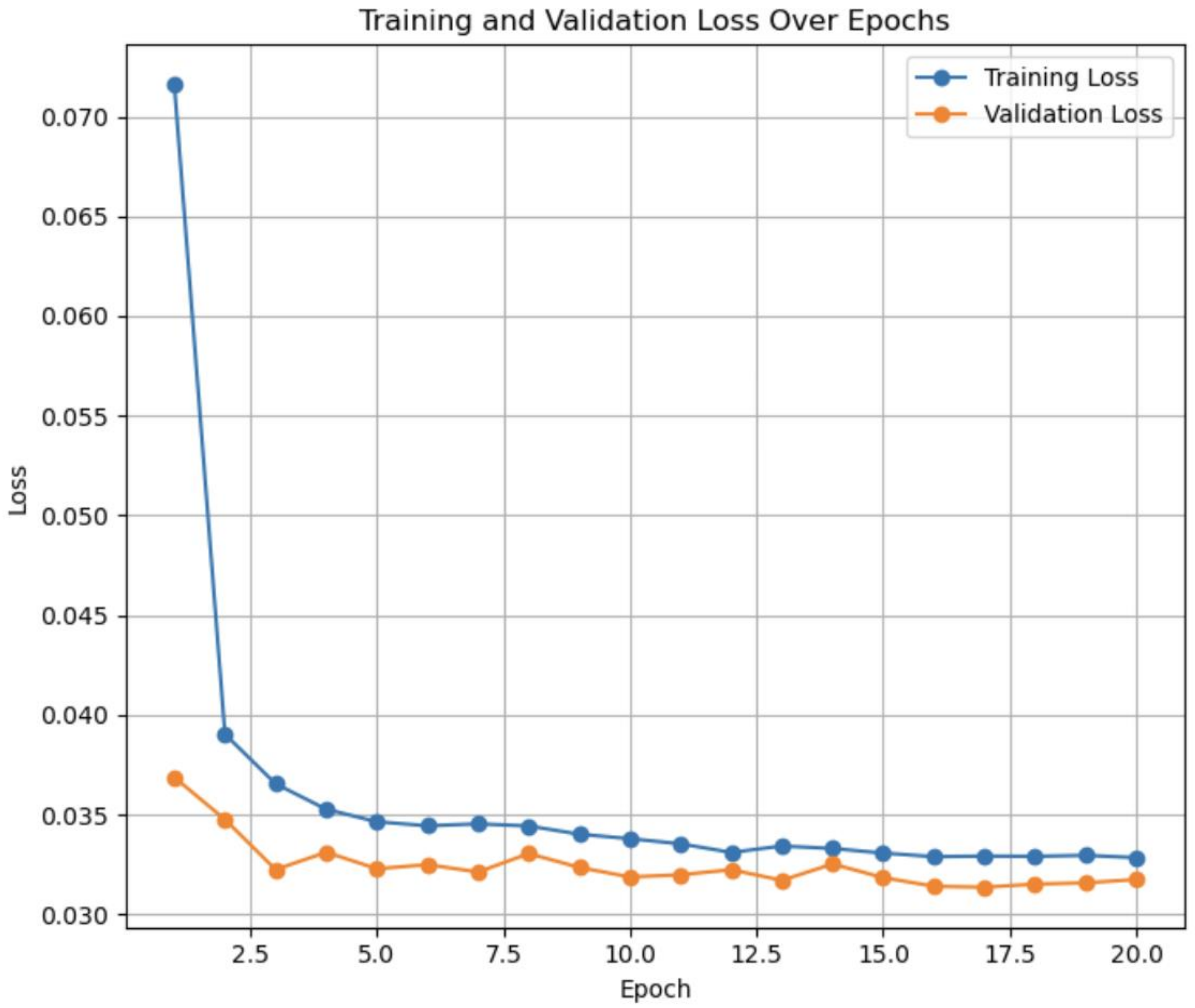
3. Model Evaluation Success Criteria

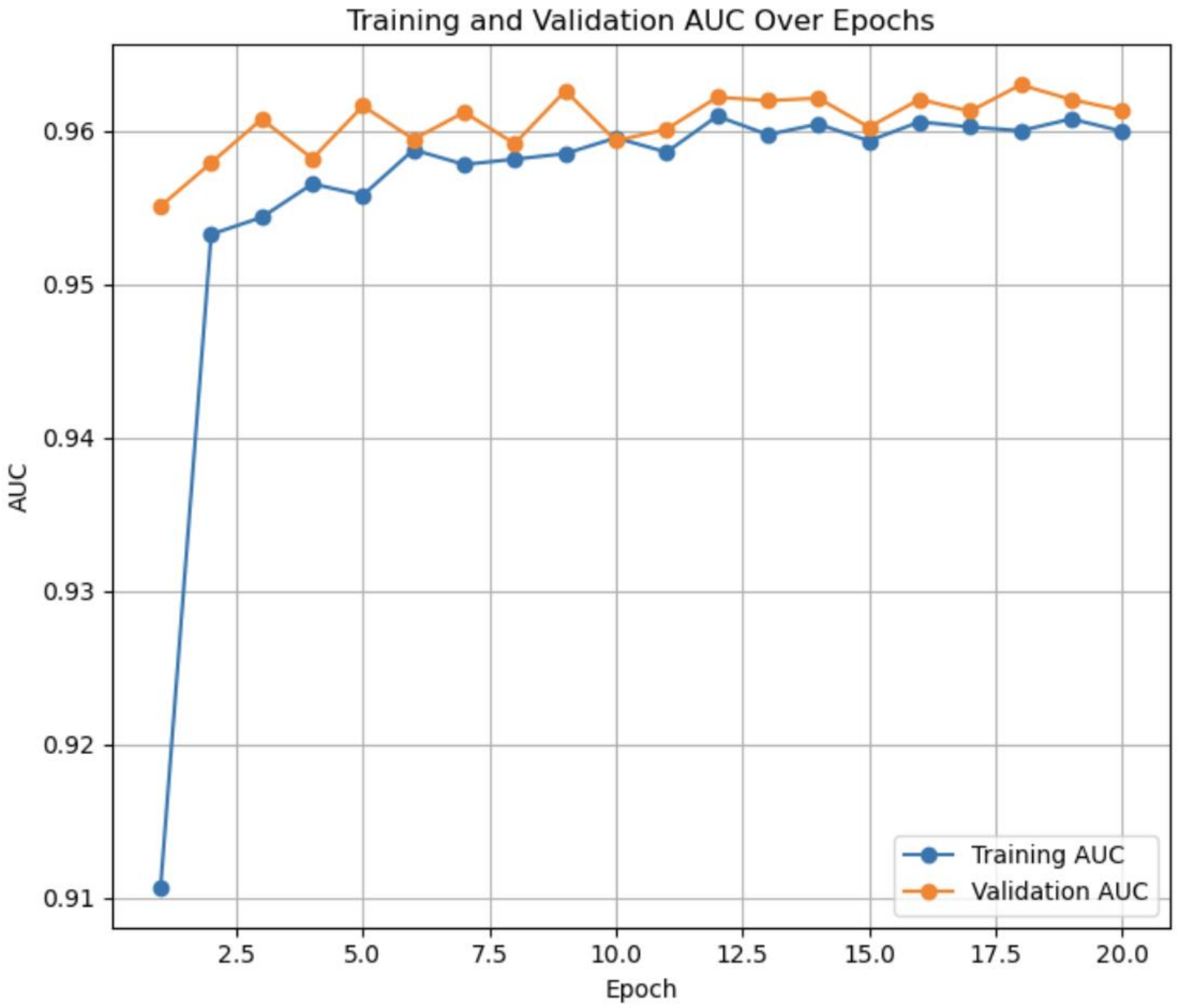
- **Random Forest**:
  - Achieve high ROC-AUC scores on both training and validation sets.
  - Compare Predicted churn probabilities for the test set vs Actual Churn Rate.
- **Neural Network**:
  - Training and validation loss decrease steadily over epochs.
  - ROC-AUC scores improved over time, indicating good generalization.
  - Compare Predicted churn probabilities for the test set vs Actual Churn Rate.

Model Training & Evaluation Results

- Inspecting at the AUC curves, neural network model performs better than random forest on both training and validation set, suggesting low bias and variance for neural network model.

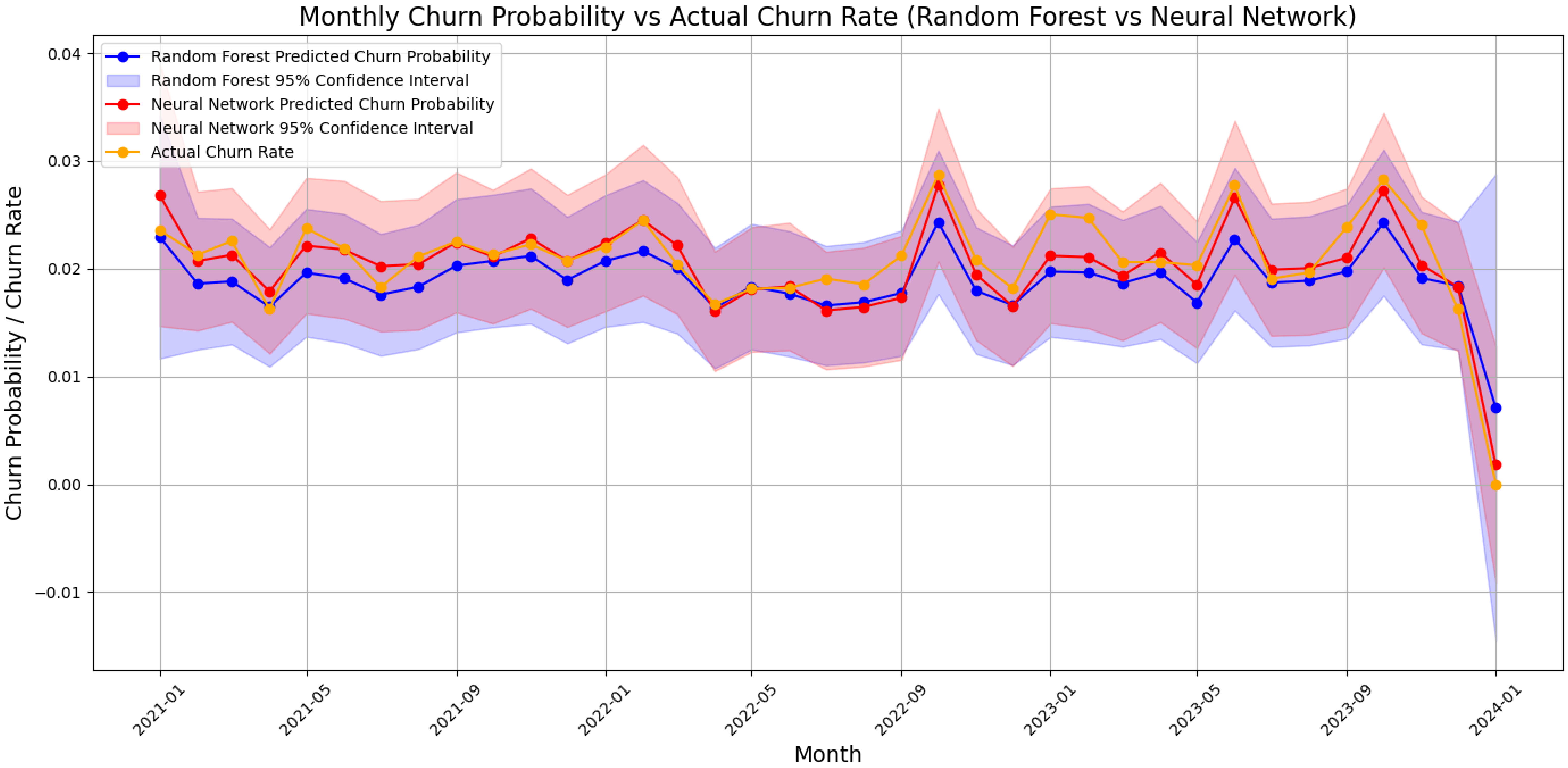






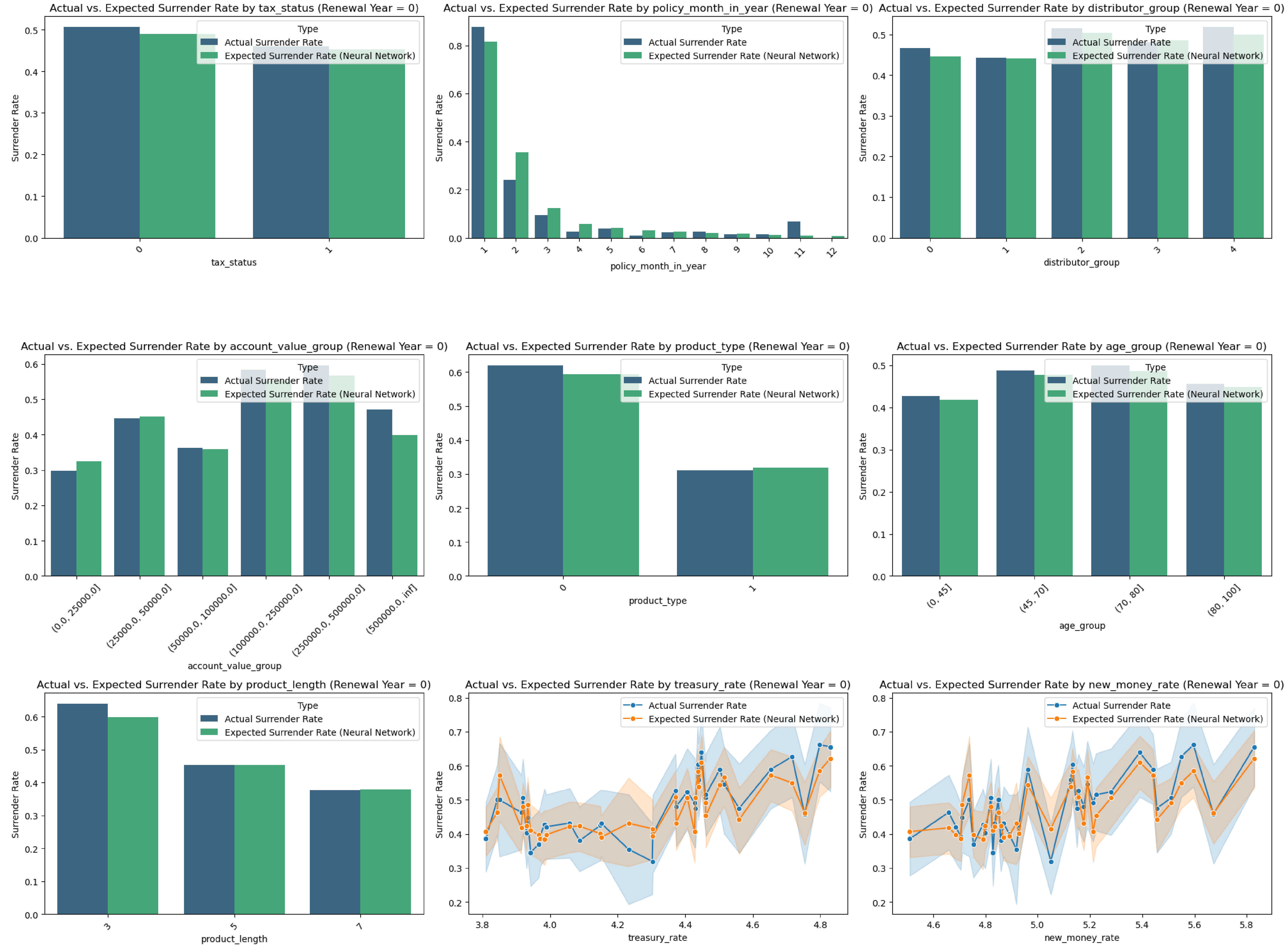
Model Training & Evaluation Results - Continued

- Inspecting at the actual vs expected churn rate, neural network model predictions are more closer to actual churn than random forest model predictions.



Period	Actual Surrender Rate	RF Predicted Surrender Rate	RF Confidence Interval	RF A/E Ratio	NN Predicted Surrender Rate	NN Confidence Interval	NN A/E Ratio
Inception-to-Date	2.14%	1.92%	(1.82%, 2.02%)	111.4%	2.07%	(1.96%, 2.17%)	103.5%
Last Twelve Months	2.24%	1.97%	(1.79%, 2.14%)	114.1%	2.12%	(1.94%, 2.30%)	105.8%

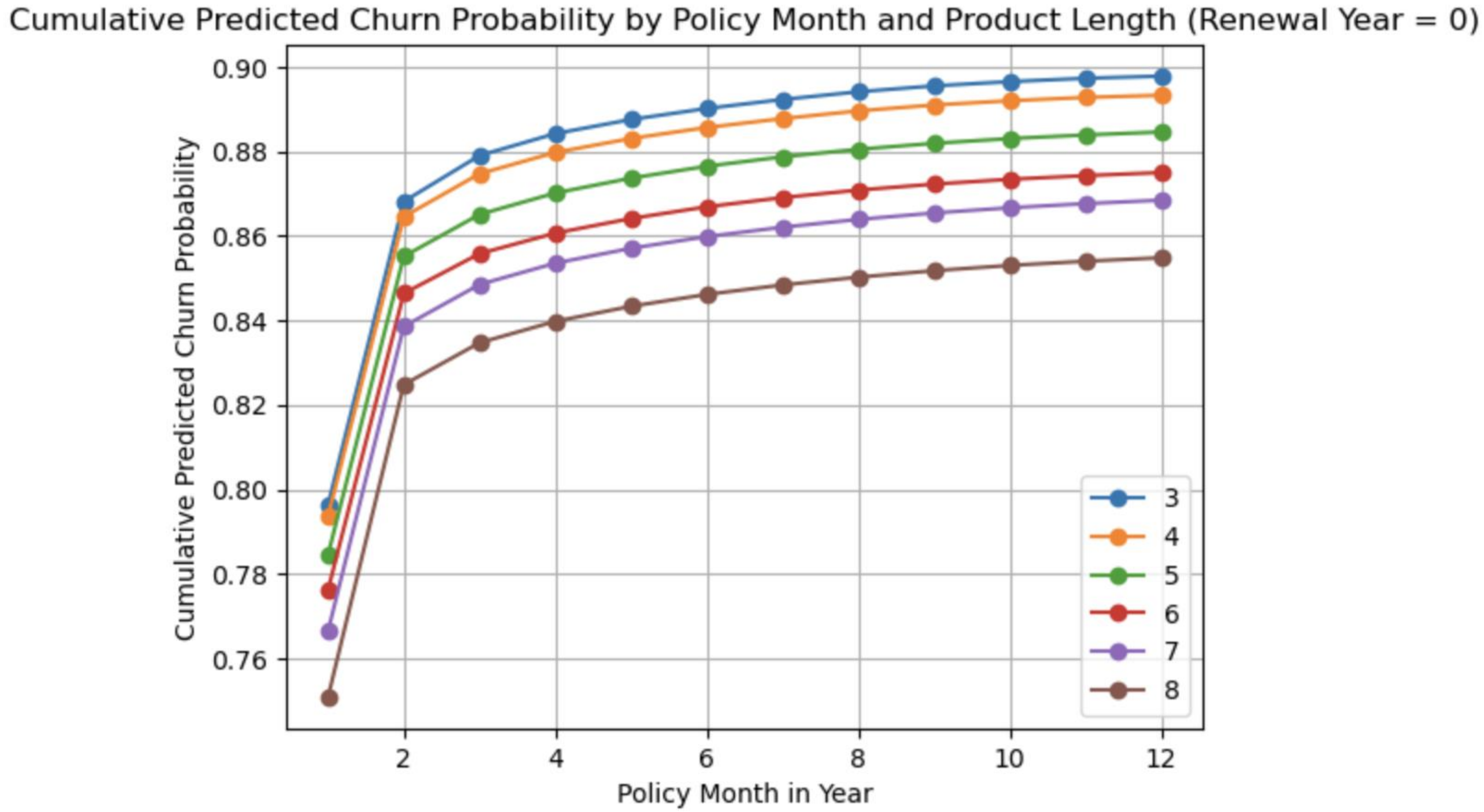
Model Training & Evaluation Results – By All IV Variables



Business Impacts

Predicted Cumulative Churn Rate by Product length – Renewal year

- This analysis will assist the pricing team in determining the optimal price for a new product with varying lengths (e.g., 4 years where we don't current sell). The predicted churn rates can be incorporated into the pricing model to set appropriate prices or crediting rates, ensuring customers are incentivized to purchase the annuity while maintaining a high Internal Rate of Return (IRR).



IRR is the discount rate that makes NPV zero:

$$0 = \sum_{n=1}^{N} \frac{CF_n}{(1+IRR)^n}$$

Where **N** = total periods, **CF_n** = cash flow at **n**, and **IRR** = Internal Rate of Return.

Competitor/ Interest Rate Sensitivity for Treasury & Investment teams

- The challenge in finding the true sensitivity for surrenders to competitor rates/ interest rates is finding good proxy to them. We don't have reliable data on our competitors and hence we use treasury rates as proxy.

- Below is a table showing rate sensitivity and its effects on churn rates. This analysis helps finance/ treasury teams understand liquidity risk in case of more churn and helps investment team make better investment/ hedging decisions.

	Cumulative Surrender Rate in First Renewal		
	Projected	-75 bps	+75 bps
Proposed	90%	84%	96%

Performance of Model by Top 10 Distributors in 1st Renewal Period– for Pricing strategies

- This shows performance of the model particularly in first renewal year for top 10 distributors.
- This table is a Proof of concept to show that using the projected rate for each distributor we can calculate the Internal Rate of Return (IRR) from pricing stand point to help company make decision regarding which distributor to focus relations with or which distributor deserves compensation increase or decrease

Distributor	Predicted Churn Rate
Broker 1	72.6%
Broker 2	76.7%
Bank	80.2%
Independent Marketing Org 1	82.6%
Independent Marketing Org 2	86.5%
Independent Marketing Org 3	88.7%
Direct Broker 1	92.7%
Direct Broker 2	94.9%