Analysis of Churn Data

What is customer churn?

Impact of Customer Churn

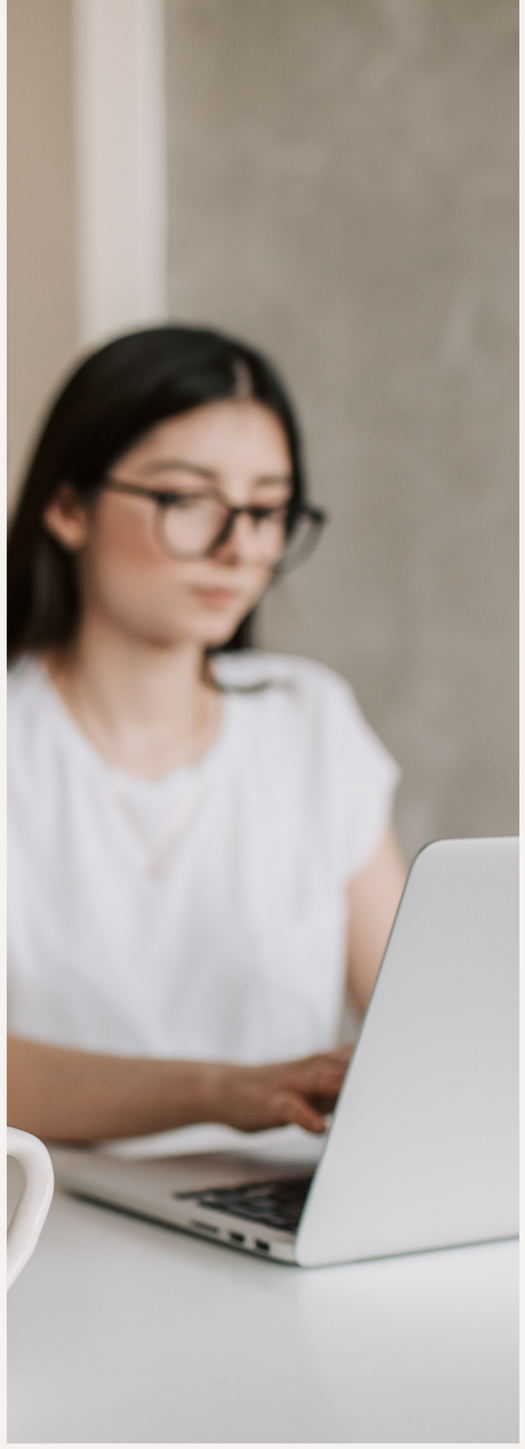
Increased Churn Rates Affect Revenue

Customer attrition, or churn, is a complex problem which can cost a business a substantial amount of revenue each year. According to Gartner, a mere 5% improvement in customer retention can increase business profits anywhere from 25% to 125%. (1)

Improving Retention Strategies for Customers

It is therefore important to implement retention strategies to decrease churn rates. Our high level analysis will highlight some of the major factors impacting your churn rate.

The Impact of Churning Customers



Customers Churn, an overview

A visual representation of customer churn data can help businesses understand why customers are leaving and remedy the situation.

Strategy, a plan

By implementing strategies such as improving customer service and creating innovative retention programs, businesses can reduce the negative effects of customer churn.

Data analysis, the core

Implementing AI technology to predict outcomes strategies can help make the right choices when customer retention programs.

We have analyzed your data and have have found many factors that are impacting customer churn.

Data Binning - Preparation for EDA

Reason for Binning

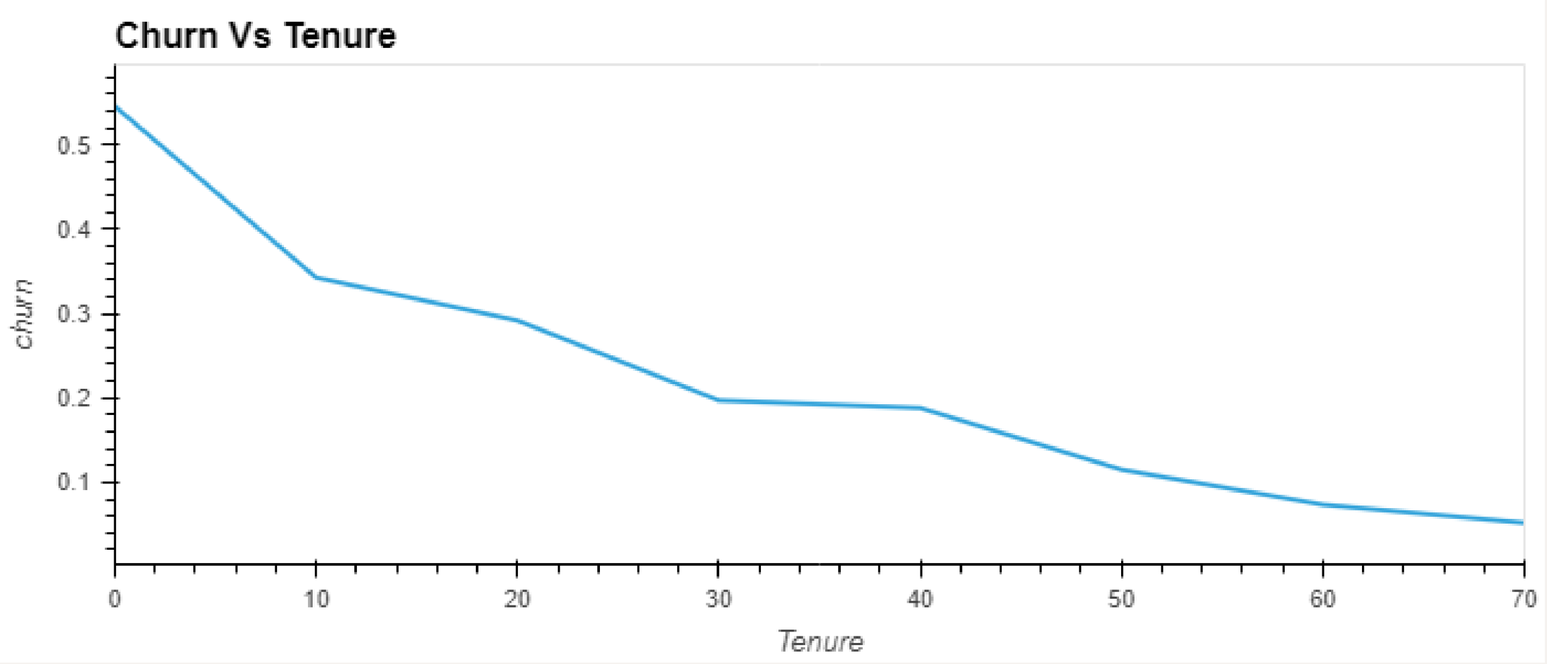
- Cannot calculate actual percent churn for individual rows. Customer churned or did not churn.
- Binned data can have many points in a bin, so an average churn % can be calculated and used for Exploratory Data Analysis
- Average % churn needed for Exploratory Data Analysis, EDA

Steps and Overview

- Generate histograms to visualize distributions of counts for the columns
- If a column had over 8 levels, create bins (applied to all floating point columns)
- If desired bins was symmetric and contiguous, used list comprehensions, otherwise bin edges were manually generated
- Functions then written to generate average churn grouping by the specific binned column.

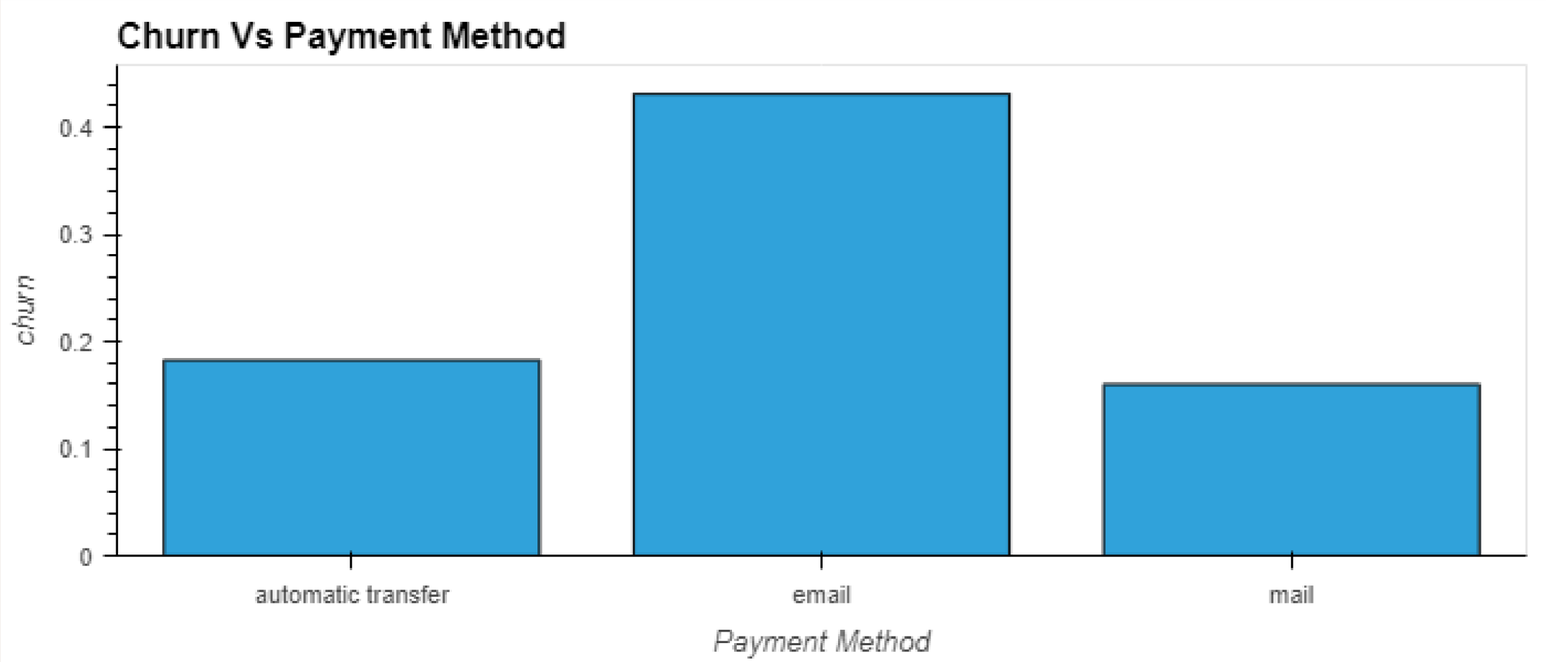
Customer Tenure

The longer a customer has had service the less likely they are to leave.



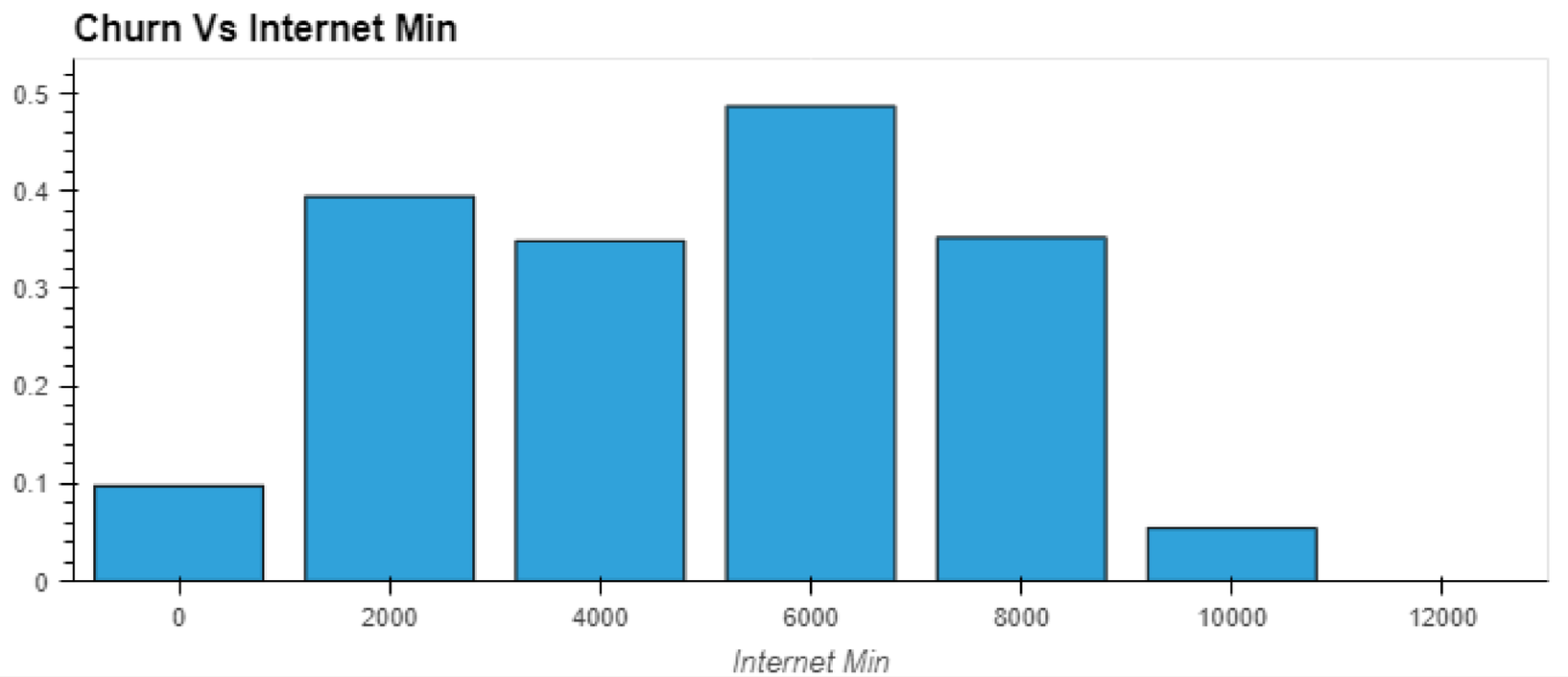
How the bill is paid

Customers payment method impacts churn.



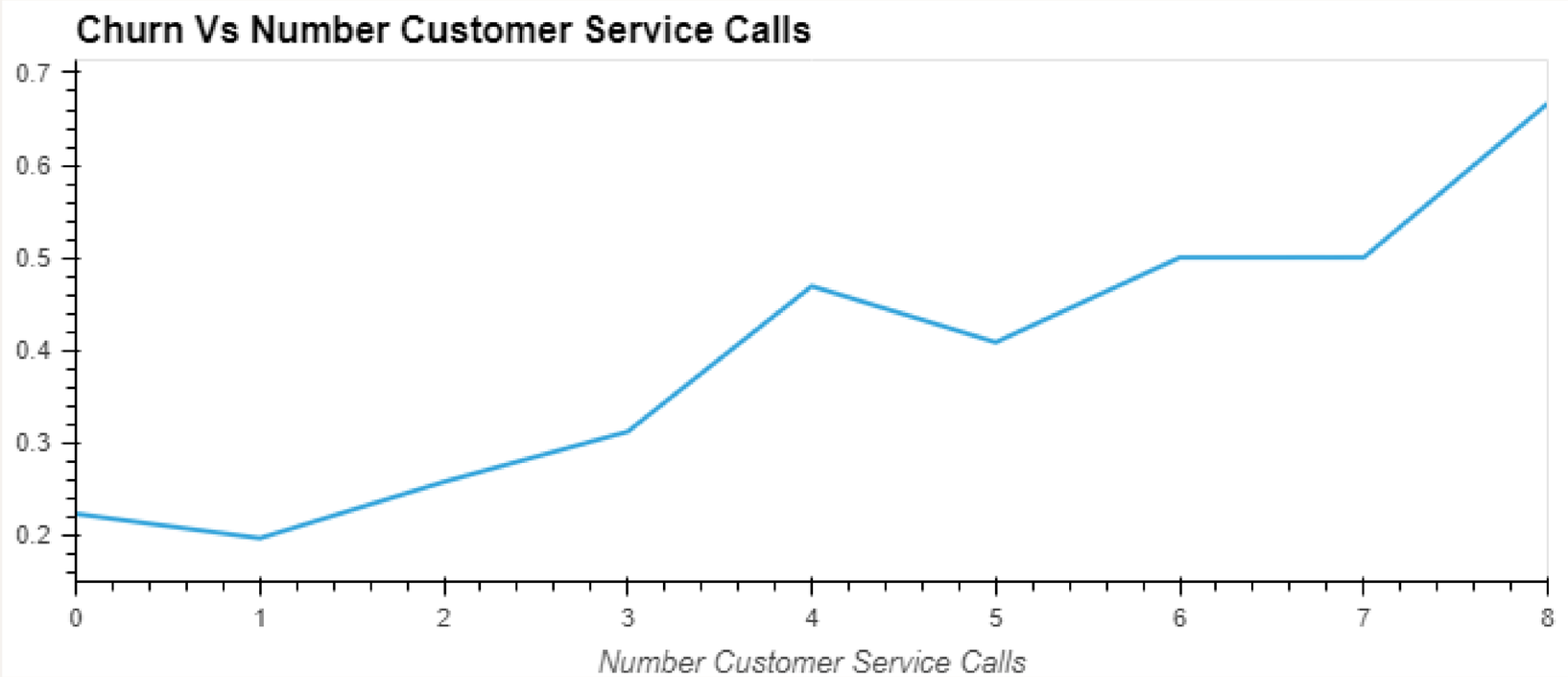
Internet Minutes Used

Customer internet usage impacts churn.



Customer Service Calls

Churn increases with more Customer Service Calls.



Interactive Chart

Click on the image below to navigate to our interactive chart.

Or navigate to: https://servicemanager.mycomputerworks.com/AI/interactive_plot.html

Customer Churn Prediction

Initial Model - Machine Learning Logistic Regression

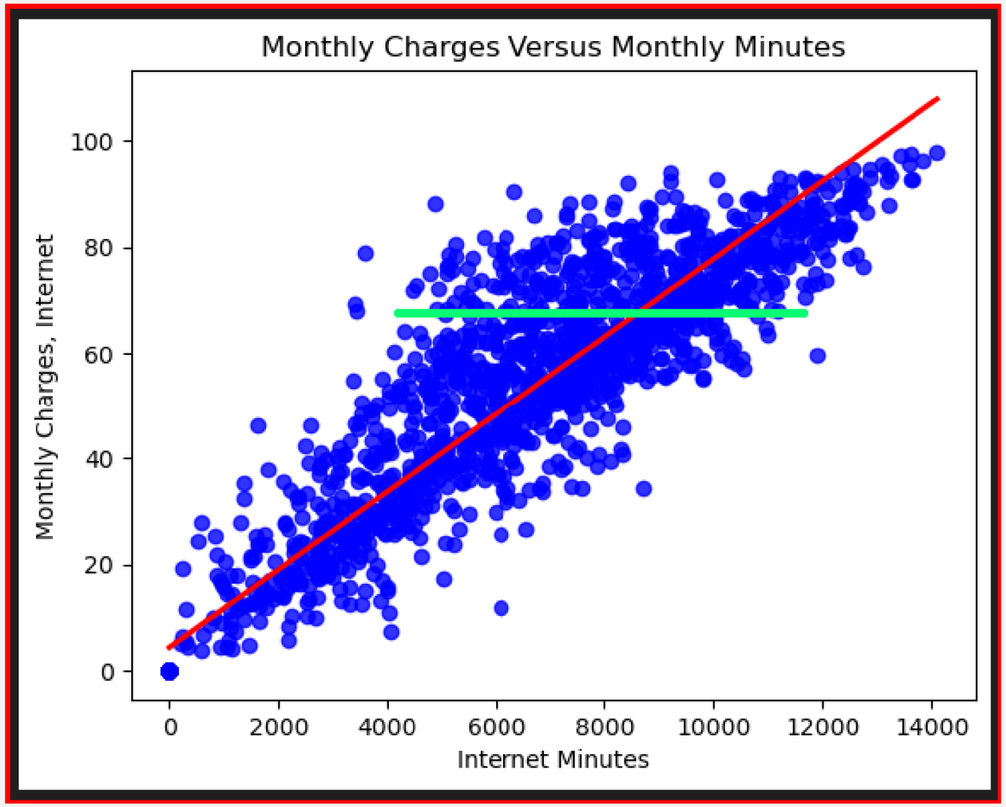
Logistic Regression Model Generation

Steps and Overview

- Using features of interest, eliminate missing data
- Encode ‘attribute’ columns to numeric
- Added internet charges per min column (and drop internet minutes per month)
- Random split of data into test versus train datasets. Test was 30% of the data.
- Define Scaling Function and apply to training data
- Train the model using scaled data (supervised machine learning)
- Apply Scaling Function to test data
- Generate predicted churns for the test data
- Evaluate performance:
  Across the test data how did predicted churn compare with actual churn?
- Generate “confusion matrix” and “feature importance” graphs

Correcting correlated columns

- Internet monthly charges correlated with Internet Minutes, but lots of spread
- Correct by computing Internet Charges per Min and dropping Internet Minutes
- Greatly improved model characteristics

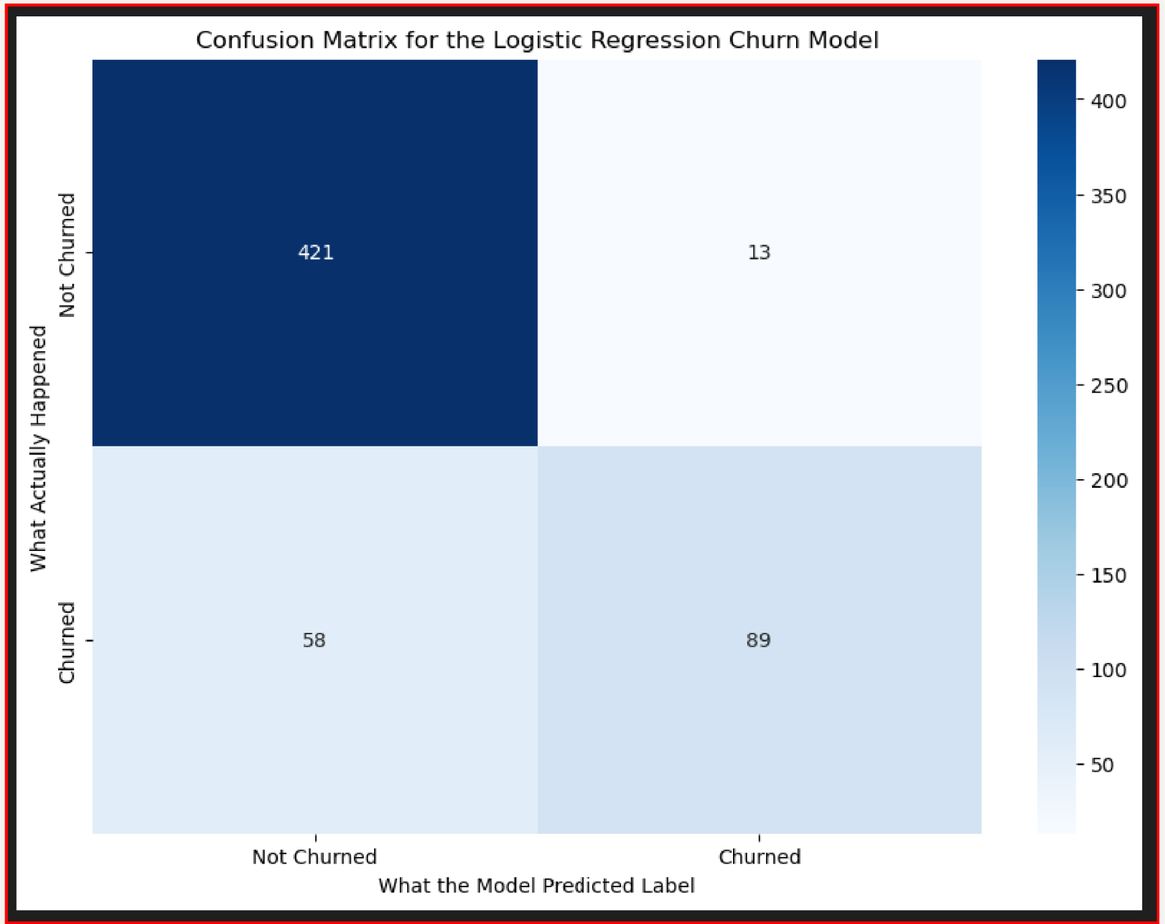


Overall Model Performance (Confusion Matrix)

Overall Model Accuracy: 87.8%

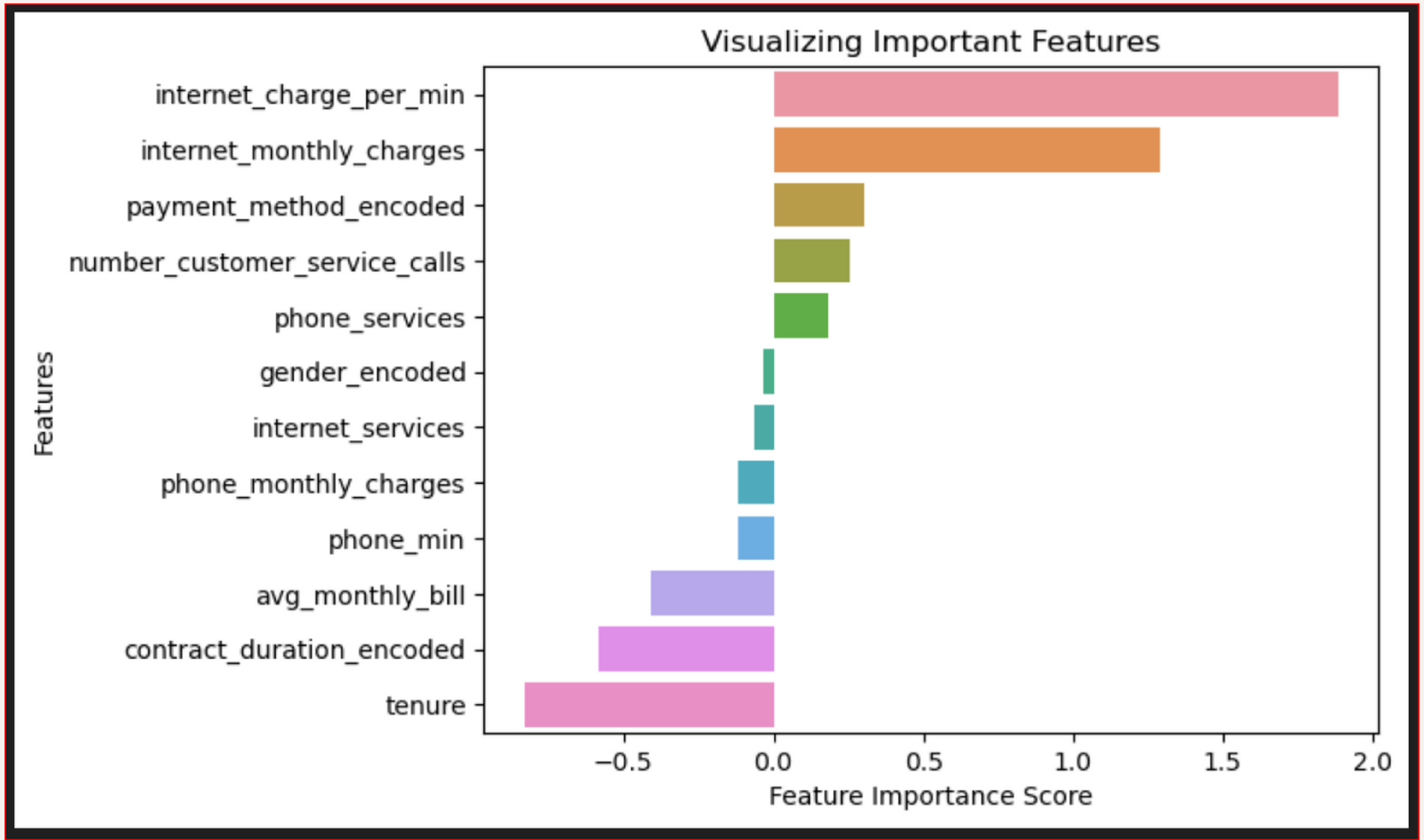
If the model predicts "No churn", it is correct 87.9 % of the time

If the model predicts "Churn", it is correct 87.3 % of the time



Feature Importance Visualized

Based on logistic linear model coefficients from standardized dataset



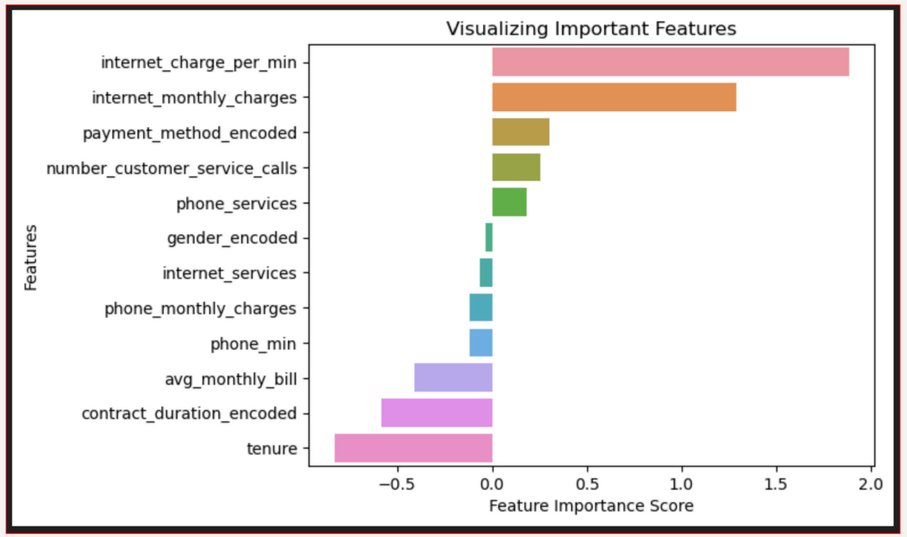
Summary from Baseline Model

Performance:

- Overall Model Accuracy: 87.8%
- Performance is about the same for predicted churn versus predicted retention
  - 87.3% correct if a churn is predicted
  - 87.9% correct if customer retension is predicted

Most important factors

- Internet charge per min (higher = more churn)
- Internet monthly charges (higher = more churn)
- Tenure (longer = less churn)
- Contract duration (longer = less churn)



Summary

Based on our analysis we have been able to show you the key factors impacting customer churn.

We have been able to use your data to give you an example of how we can use AI to provide insight in your data.

With your permission we will provide a proposal for Phase 2 where we will leverage our AI skills to provide further insight into your data and help you create processes that will lower your customer churn and increase your profit.

References

1) Gartner Group and “Leading on the Edge of Chaos”, E. Murphy and M. Murphy
2) ChatGPT used for ideas and code generation for the machine language model
3) “Closest” function which returns the closest bin range given the actual value is from https://www.geeksforgeeks.org/python-find-closest-number-to-k-in-given-list/
4) Data from GitHub Repo: griddynamics/ rnd-gcp-starter-kits https://github.com/griddynamics/rnd-gcp-starter-kits/tree/main
5) Churn analytics in the technology and telecom industries using Google Vertex AI (https://blog.griddynamics.com/churn-analytics-in-the-technology-and-telecom-industries-using-google-vertex-ai-a-reference-notebook)