Customer Churn Analysis for Telecommunications Industry

Data-Driven Insights and Retention Strategies through K-Means Clustering & Keras ANN

This project builds a data-driven customer churn prediction model for a telecommunications company, leveraging feature engineering, K-Means clustering, and Keras-TensorFlow ANN to identify high-risk customers and recommend targeted retention strategies that reduce churn and increase customer lifetime value.

Prepared by:
Group 2, Work Integrated Learning - Data Analytics - WILDA25_06
Yake Liu– Project Manager,
Tianyi Yin– Data Engineer - Preprocessing & Feature Engineering,
Lingxin Zhou– Data Analyst - Clustering Analysis,
Yu-chieh Hsu– Data Analyst - Predictive Modeling,
Xinyue Li– Business Analyst

21/09/2025

Contents

1. Introduction

Customer churn is still a big problem for the telecommunications industry because it hurts profits and long-term growth. The goal of this project is to analyse customer data and come up with useful insights that can help with retention strategies. To achieve this, we used two different methods that worked well together: k-means clustering to group customers based on tenure and monthly charges, and an artificial neural network (ANN) based on Keras to predict the probability at the individual level.

By combining these two methods, we get a full picture of behaviour, find groups that are at risk, and suggest targeted actions that balance retention, revenue growth, and customer satisfaction. This report outlines our key findings, supporting evidence, and strategic recommendations for stakeholders.

2. Summary of Key Findings

(1) Key findings from customer segmentation

We effectively segmented 7,043 customers into four groups based on tenure and MonthlyCharges using K-means clustering. The Elbow Method and the Silhouette Score were employed to assess the segmentation. Ultimately, we determined that four clusters were the most optimal compromise between interpretability and model fit.

Customer Segment Profiles

Cluster 0 (31.3%) — New + High Price, churn (49.2%) → largest and highest-risk segment.

Characteristics: New customers on high-priced plans; low stability—nearly half are likely to churn.

Cluster 1 (24.6%) — New + Low Price, medium churn (24.5%) → growth potential.

Characteristics: New customers with lower spend; mid-level churn risk.

Cluster 2 (27.7%) — Long-tenure + High Price, lower churn (15.7%) → high-value customers.

Characteristics: Long-standing customers with higher spend; relatively low churn.

Cluster 3 (16.5%) — Long-tenure + Low Price, very low churn (4.8%) → stable, loyal segment.

Characteristics: Long-standing customers with lower spend; highly stable.

Implication: The four segments clearly show different risk and potential. The company should adopt **differentiated retention and value-growth strategies**:

- Reduce churn in Cluster 0,
- Develop growth potential in Cluster 1,
- Reinforce high value in Cluster 2,
- Maintain stability and modestly uplift Cluster 3.

(2) Key findings from churn prediction analysis:

- **Learning health:** Train/val AUPRC up, loss down, curves aligned; bell-shaped weight distribution → minimal overfitting.

- **Imbalance handling:** Churn ≈ 27:73. Threshold chosen via PR curve; F1-optimal recommended (Accuracy-optimal as a cost-sensitive backup).
- **Optional baseline:** Calibrated Gradient Boosting: slightly lower overall accuracy but better minority-class (churner) handling.

3. Identification of Factors Contributing to Churn and Retention:

(1) Tenure & Contract Length
Finding: In all visuals, **longer tenure → lower churn**. One- and two-year contracts exhibit a tenure commitment and consistently lower attrition rates; month-to-month contracts exhibit the highest churn. This is consistent with clustering: segments that are new (C0, C1) are more risky than long-tenure segments(C2, C3).
Interpretation: The likelihood of volatility is diminished by relationship maturity and commitment.

(2) Price / MonthlyCharges (with segments)
Finding: MonthlyCharges is the **second-most critical feature** (Figure 10). The highest observed attrition (~49%) is observed in the new + high price (C0) segment, while the long-tenure + high price (C2) segment is significantly lower.
Interpretation: Retention is not guaranteed by price alone; it is contingent upon tenure and commitment. High-spenders tend to fluctuate during their initial tenures, but they tend to remain consistent over time.

(3) SeniorCitizen
Finding: Heatmaps (Figure 11) show **seniors have higher churn** than non-seniors **at every tenure level**.
Interpretation: Age-related factors (e.g., affordability/usability) keep churn elevated even as tenure grows.

(4) Dependents
Finding: Customers **with dependents** have **lower churn,** particularly **during short tenures** (shown by cooler bands in Figure 11).
Interpretation: Domestic responsibilities are associated with more consistent consumption and less early churn.

(5) Internet Service Type (with Contract)
Finding: At the same contract length, **fiber-optic users churn slightly more** than **DSL**, with the gap most visible for **month-to-month**. (Shows in Figure 12)
Interpretation: Service type interacts with churn risk; fiber skews riskier under flexible contracts.

(6) Class Imbalance & Operating Point
Finding: With an imbalanced dataset (~27:73), the model was evaluated with **AUPRC** alongside accuracy (test accuracy ~79%, AUPRC ~0.60). Thresholds were chosen from the **precision–recall curve** (F1-optimal vs accuracy-optimal).
Interpretation: Use the **F1-optimal threshold** by default to balance precision/recall on churners; accuracy-optimal is a cost-sensitive fallback.

4. Recommendations for Targeted Retention Strategies:

(1) Let's review the Customer Segment Profiles in Section 2, based on k-means clustering analysis, we got the Strategies:

Cluster 0 (31.3%) — New + High Price, churn 49.2% → largest and highest-risk segment.
Strategy: Focus segment, onboarding support, personalised discounts, rapid issue resolution.
Cluster 1 (24.6%) — New + Low Price, medium churn (24.5%) → growth potential.
Strategy: Growth segment, offer small value-added services and promote package upgrades.
Cluster 2 (27.7%) — Long-tenure + High Price, lower churn (15.7%) → high-value customers.
Strategy: High-value segment, invest in VIP programs, proactive satisfaction checks, premium service perks.
Cluster 3 (16.5%) — Long-tenure + Low Price, very low churn (4.8%) → stable, loyal segment.
Strategy: Loyal segment, maintain satisfaction while exploring cross-selling opportunities to increase lifetime value.

(2) Based on keras ANN analysis, we got the Strategies:

- **Thresholds aligned to business goals**
  Use the precision–recall curve to pick an operating point:
  Accuracy-optimal: capture more potential churners (fewer unnecessary contacts).
  F1-optimal: minimize false positives (better balance of precision/recall).
  This flexibility lets teams choose **cost control vs. maximum retention.** *(For imbalanced churn data, F1-optimal is the recommended default.)*
- **Improve revenue and customer experience**
  Apply **targeted incentives, dedicated support,** and **personalized outreach** to high-risk customers to raise satisfaction and loyalty**.**
- **Enable granular, data-driven operations**
  Use the model's probability to **tier customers** (high / medium / low risk) and run **differentiated playbooks**: intensive save offers for high risk, light nudges for medium risk, and experience maintenance with upsell testing for low risk.

5. Documentation of Limitations and Proposed Solutions:

(1) Class imbalance (churners ≈ 27%, non-churners ≈ 73%)

- **Problem:** The model might overlook Churners because they are a minority.
- **Fix:** Use segment-specific thresholds (e.g., a lower cutoff for Cluster 0), track PR-AUC, default to the F1-optimal threshold, and use class weights.

(2) Limited features (due to source data)

- **Problem:** Predictive lift is limited because behavioral and time-window fields are absent from our initial database.
- **Fix (with better data):** extend data collection to include payment history, support tickets/complaints, and usage for 30/60/90 days. After those histories are available, create rolling characteristics (volatility, deltas, and levels).

(3) Tuning time/cost

- **Problem:** Hyperband trials can be slow and block iteration.

- **Fix:** Cap trials/epochs, keep EarlyStopping on, use GPU/mixed precision, schedule runs off-peak, and when rushed, reuse last good hyperparameters.

(4) Explainability & trust

- **Problem**: To many people, the ANN model looks like a "black box" — they can't see how it makes decisions.
- **Fix**: Give simple reasons for why each customer group is at risk, highlight the main factors that influence predictions, and share a one-page summary showing what data was used, how well the model works, its limits, and when it might need updates.

(5) Privacy & outreach compliance

- **Problem**: Reaching out to customers too often, or in the wrong way, can hurt satisfaction and lead to complaints.
- **Fix**: Only collect the info you truly need, protect personal data, limit how many times you contact or discount customers, and keep track of feedback. If issues show up, scale back quickly.

6. Present Results and Recommendations to Stakeholders

Our analysis combined **k-means clustering** and **a neural network churn prediction model keras**. Together, they highlight clear customer patterns and provide actionable strategies to reduce churn and grow value.

Key Results:

- **Cluster Insights:** Four distinct customer groups were identified, each showing different churn risks and value contributions.
- **ANN Insights:** The model enables **early identification of high-risk customers**, allowing proactive retention.
- **Common Finding:** Both methods confirm that new high-price users are the most vulnerable to churn, while **long-tenure, high-value customers** drive sustainable revenue.

Final strategies:

- **Early Identification and Retention of High-Risk Customers**

Rationale: Both clustering and ANN show new high-price users are most likely to churn.
Actions: Implement a churn-risk alert system that flags customers with high predicted probability. Offer targeted retention incentives, such as one-month fee waivers or free plan upgrades.

- **Enhance Onboarding for New Customers**

Rationale: New users, regardless of price tier, face higher churn risk in early months.
Actions: Launch a "Welcome Program" with step-by-step setup guides and 24/7 chatbot support. Provide dedicated customer service for the first 3 months to resolve early issues quickly.

- **Deliver Premium Care to High-Value Customers**

Rationale: Long-tenure, high-price users are stable but critical for profitability.
Actions: Create a VIP membership program with loyalty points, priority service, and exclusive perks. Conduct quarterly check-ins by account managers to ensure satisfaction and capture feedback.

- **Unlock Value from Loyal Low-Price Customers**

Rationale: Long-tenure, low-price users rarely churn but have untapped growth potential.
Actions: Promote cross-selling opportunities, e.g., bundling internet or streaming services. Offer tiered upgrade discounts, such as 20% off after 12 months of continued service.

7. Visualization

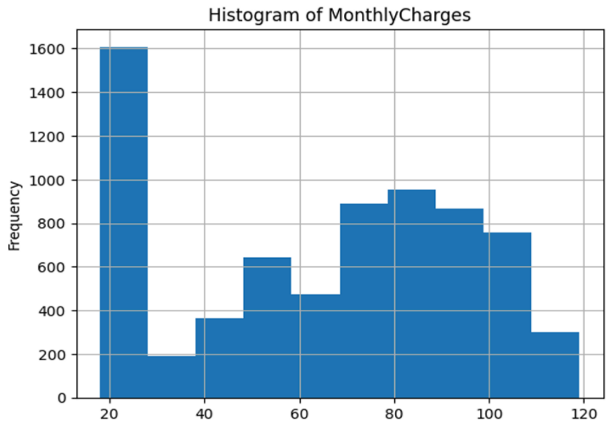

Figure 1

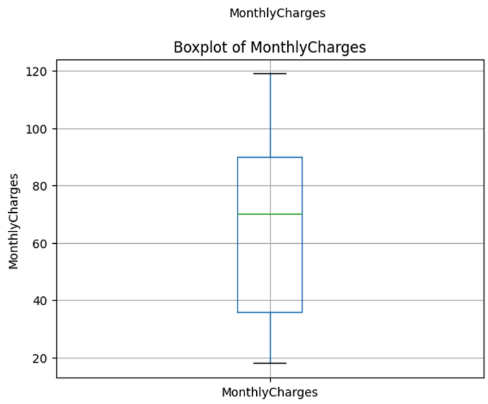

Figure 2

Figure 1 — Histogram of MonthlyCharges

- Most customers pay around $20–30, and there's another large group around $70–90.
- Fewer customers pay $30–50.
- Very few pay above $100.

Figure 2 — Boxplot of MonthlyCharges

- Most bills fall roughly between $40 and $90.
- Only a small number are below $30 or above $100.

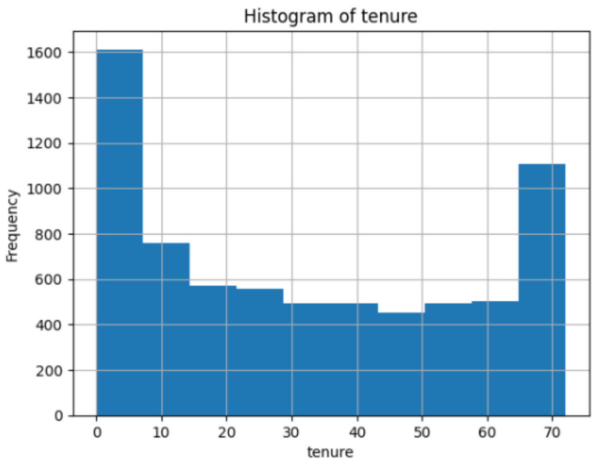

Figure 3

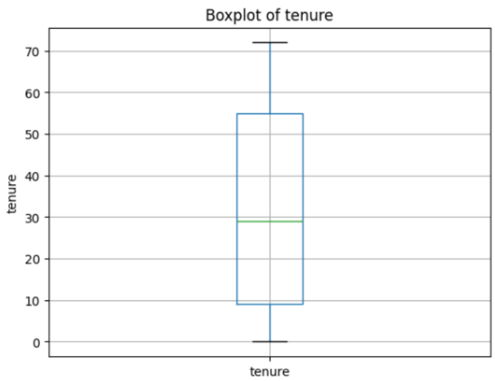

Figure 4

Figure 3 — Histogram of tenure

- Most customers are either very new (0–6 months) or very long-tenure (60–72 months).
- Fewer customers sit in the 10–50 months range.

Figure 4 — Boxplot of tenure

- Most tenures are around 10–55 months.
- Only a small group are brand new or over 60 months.

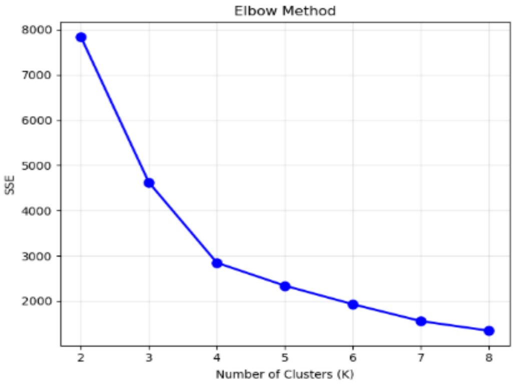

Figure 5

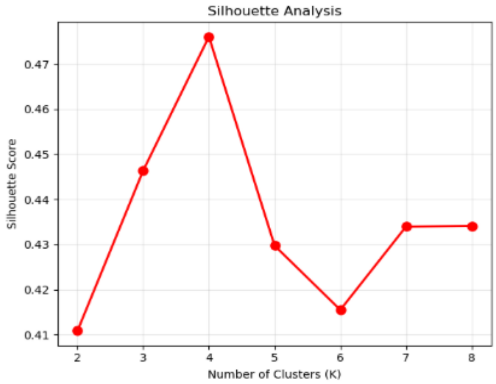

Figure 6

Figure 5 — Elbow Method

- As we increase the number of clusters from K=2 → 4, the error drops a lot.
- After K=4, the curve flattens — adding more clusters only gives small improvements.
- Takeaway: 4 clusters is the sensible cut-off (good fit without over-splitting).

Figure 6 — Silhouette Analysis

- The score is highest at K=4 (best separation).
- K=3 is decent, K=5 is lower, K=6 dips the most, and K=7–8 recover slightly but don't beat K=4.
- Takeaway: The quality of clustering is best with 4 clusters.

Overall: Both charts point to **K = 4** as the right number of customer segments.

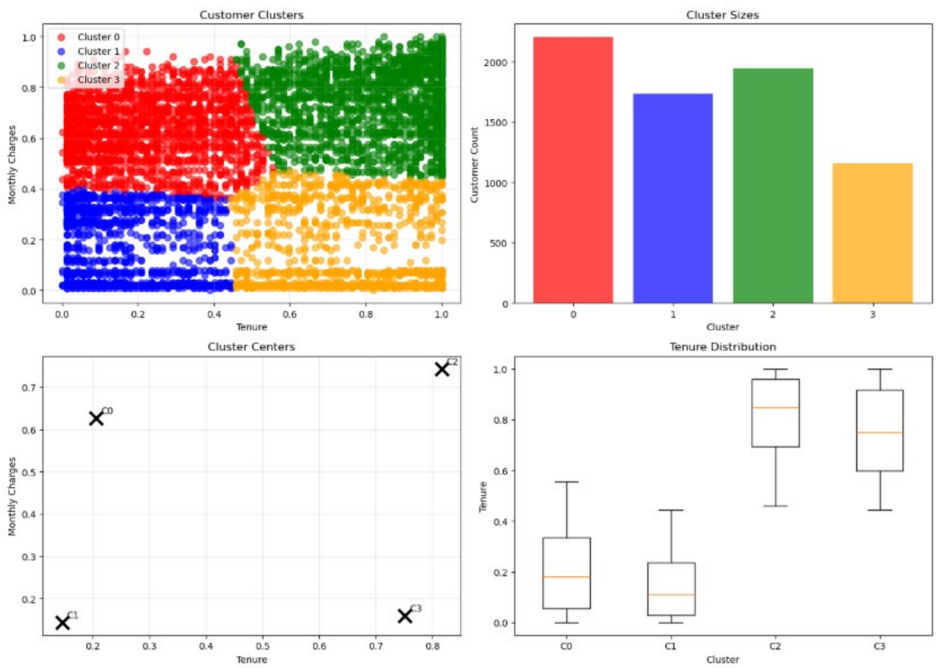

Figure 7

Top-left — Customer Clusters (scatter of Tenure × MonthlyCharges)

- Cluster 0 (red): Low tenure + high monthly charges.
- Cluster 1 (blue): Low tenure + low monthly charges.
- Cluster 2 (green): High tenure + high monthly charges.
- Cluster 3 (orange): High tenure + low monthly charges.

Top-right — Cluster Sizes (bar chart)

- Largest: Cluster 0 (red).
- Next: Cluster 2 (green), then Cluster 1 (blue).
- Smallest: Cluster 3 (orange).

Bottom-left — Cluster Centers (tenure, monthly charges)

- C0 center: Low tenure, high charges.
- C1 center: Low tenure, low charges.
- C2 center: High tenure, high charges.
- C3 center: High tenure, low charges.

Bottom-right — Tenure Distribution by Cluster (boxplots)

- C0 & C1: Tenure is mostly low.
- C2 & C3: Tenure is mostly high (upper end of the scale).

Model: "sequential"

Layer (type)	Output Shape	Param #
dense (Dense)	(None, 32)	352
batch_normalization (BatchNormalization)	(None, 32)	128
activation (Activation)	(None, 32)	0
dropout (Dropout)	(None, 32)	0
dense_1 (Dense)	(None, 64)	2,048
batch_normalization_1 (BatchNormalization)	(None, 64)	256
activation_1 (Activation)	(None, 64)	0
dense_2 (Dense)	(None, 1)	65

Total params: 8,165 (31.90 KB)
Trainable params: 2,657 (10.38 KB)
Non-trainable params: 192 (768.00 B)
Optimizer params: 5,316 (20.77 KB)

Figure 8: The model summary implemented in Keras (TensorFlow backend).

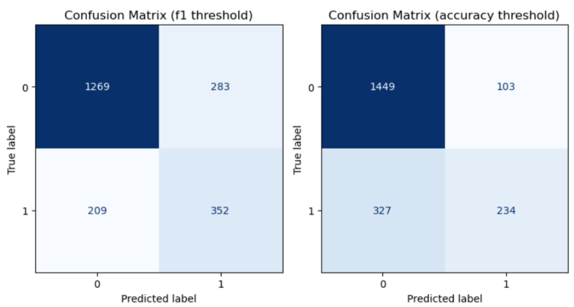

Figure 9

Confusion matrix calculated based on different thresholds: F1 score (left) and accuracy/default (right). The left matrix balances precision and recall more evenly, resulting in a higher number of correctly predicted labels overall.

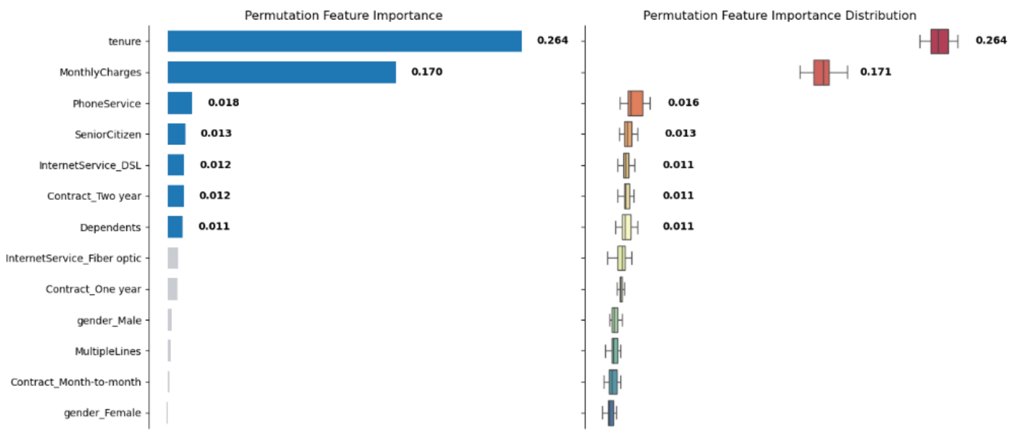

Figure 10

Permutation Feature Importance

- Tenure is the top driver by a clear margin; MonthlyCharges is second.
- Other factors (PhoneService, SeniorCitizen, DSL, Two-year contract, Dependents, etc.) matter but much less.
  → This ranking backs your takeaways: tenure dominates, price matters, and seniors/dependents still play a role.

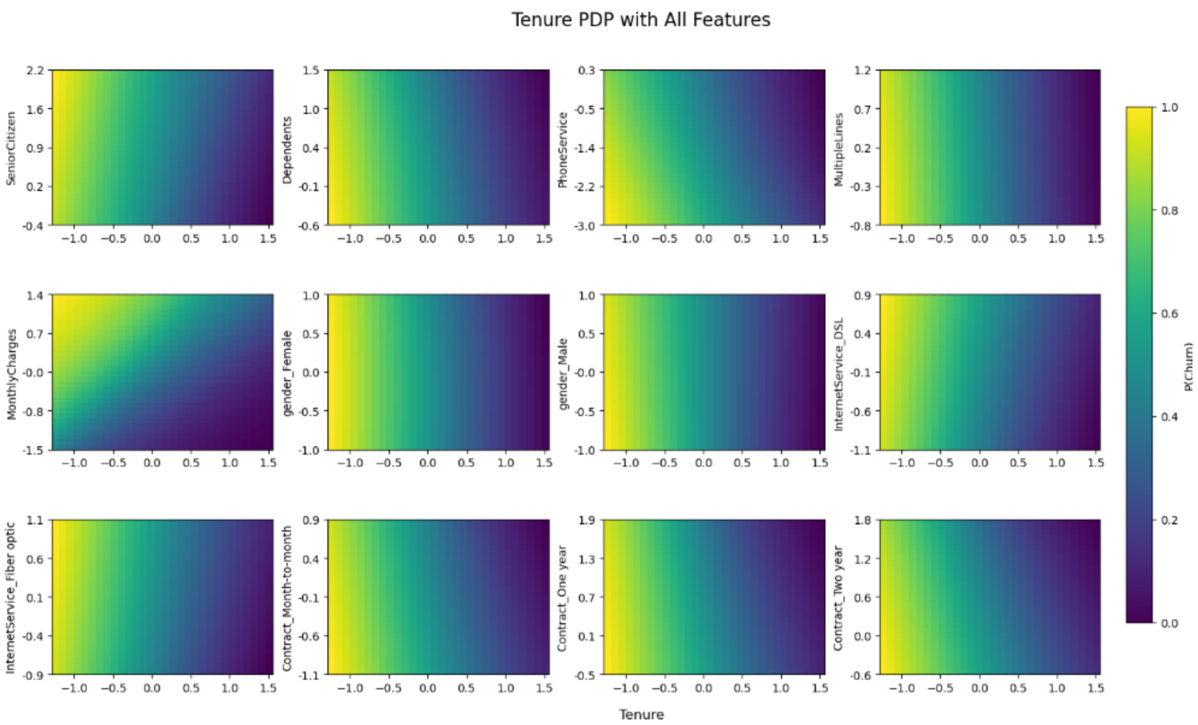

Figure 11

Tenure PDP heatmaps

- Moving right (higher tenure) the color cools → churn probability goes down as tenure increases.
- SeniorCitizen pane: the "senior = 1" band stays warmer than "non-senior = 0" at every tenure → seniors churn more at all tenure levels.
- Dependents pane: the "dependents = 1" band is cooler, especially on the left (short tenure) → customers with dependents churn less, most clearly early on.
- Contract panes: "month-to-month" stays warmer than one-year / two-year → short contracts are riskier.
- InternetService panes: fiber optic looks warmer than DSL → fiber users churn a bit more.

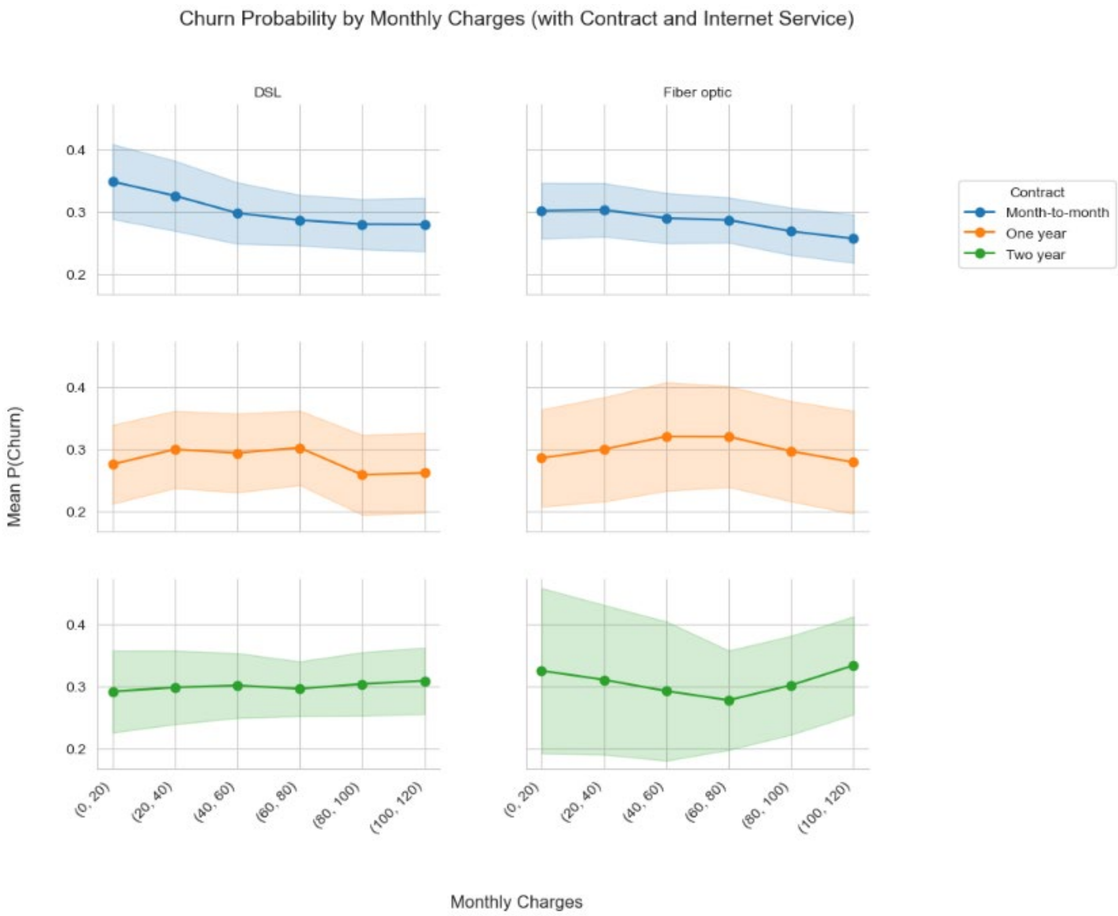

Figure 12

Churn vs Monthly Charges × (Contract, Internet)

- Month-to-month (blue) has the highest churn across nearly all charge bands and both internet types.
- One-year (orange) and Two-year (green) are consistently lower and more stable.
- At the same contract length, DSL (left column) tends to be slightly lower churn than fiber (right column); the gap is most visible for month-to-month.
→ Contract length behaves like a commitment/tenure proxy; service type also interacts with risk.

8. Conclusion

The analysis confirms that customer churn is influenced by multiple factors, including tenure, contract type, pricing, demographics, and service type. Clustering analysis revealed **four distinct customer segments** with different risk levels, while the ANN model enabled **early detection of churn-prone customers**. Together, these approaches reinforce the need for differentiated retention strategies:

- **Proactively manage high-risk new customers** with tailored incentives and onboarding support.
- **Invest in high-value long-tenure customers** through premium programs and proactive engagement.
- **Maintain and grow loyal low-price customers** with cross-selling and upgrade opportunities.

While data limitations and model complexity present challenges, the integration of segmentation and predictive analytics provides a robust framework for churn reduction. Implementing these recommendations will not only lower churn but also enhance customer experience and long-term profitability.

Reference:

https://acslearn.instructure.com/courses/1541/pages/page-s1w1p2-project-scenario-roles-and-dataset?module_item_id=147997 (Dataset source)

Géron, A. (2019) *Hands-on machine learning with Scikit-Learn, Keras, and TensorFlow*. 2nd edn. Sebastopol, CA: O'Reilly Media. Available at: http://14.139.161.31/OddSem-0822-1122/Hands-On_Machine_Learning_with_Scikit-Learn-Keras-and-TensorFlow-2nd-Edition-Aurelien-Geron.pdf (Accessed: 21 September 2025).

Amit (2024) 'Hyperband hyperparameter optimization', *Medium*, 5 December. Available at: https://medium.com/@heyamit10/hyperband-hyperparameter-optimization-d7bd66faa8e8 (Accessed: 19 September 2025).

Warnes Z. (2021) 'TensorBoard in 5 minutes', *Towards Data Science*, 5 November. Available at: https://towardsdatascience.com/tensorboard-in-5-minutes-71c5715a10d3 (Accessed: 21 September 2025).

Dobilas S. (2021) 'Gradient boosted trees for classification: one of the best machine learning algorithms', *Towards Data Science*, 7 March. Available at: https://towardsdatascience.com/gradient-boosted-trees-for-classification-one-of-the-best-machine-learning-algorithms-35245dab03f2 (Accessed: 21 September 2025).

Raskoshik V. (2023) 'Feature importance: how not fool yourself', *Kaggle*. Available at: https://www.kaggle.com/code/raskoshik/feature-importance-how-not-fool-yourself (Accessed: 21 September 2025).