b. Identification of Factors Contributing to Churn and Retention:

- Include an analysis document identifying and analysing factors contributing to customer churn and retention based on the ANN model's predictions.

(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.