a. Summary of Key Findings:

Provide a document summarising the key findings from customer segmentation and churn prediction analysis conducted using the ANN model.

(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.