Neural Network Model

Simple data processing

- **Numerical Transformation:** Converted predictors into numeric matrices
  - yes → 1; no → 0
  - Ensures compatibility with Keras tensor operations

- **Data Splitting:**
  - **Baseline model:** 80% training set(80% training + 20% validation) + 20% test set
  - **Cross validation:** 5 folds for robustness, each fold trained/validated separately

Neural Network Model

Baseline Model (Single Training: Keras Model)

- **Input** → Dense(64, ReLU) → Dropout(0.4)
  Dense(32, ReLU) → Dropout(0.4)
- **Output**: Dense(1, Sigmoid)

- **Loss Function**: Focal Loss (α = 0.8, γ = 2.0)
  - (enhancing model focus on minority class)

- **Metrics**: MCC, AUC, Precision & Recall

- **Classification Threshold:** Select the classification threshold corresponding to the optimal MCC value

Neural Network Model

Baseline Model (Single Training: Results)

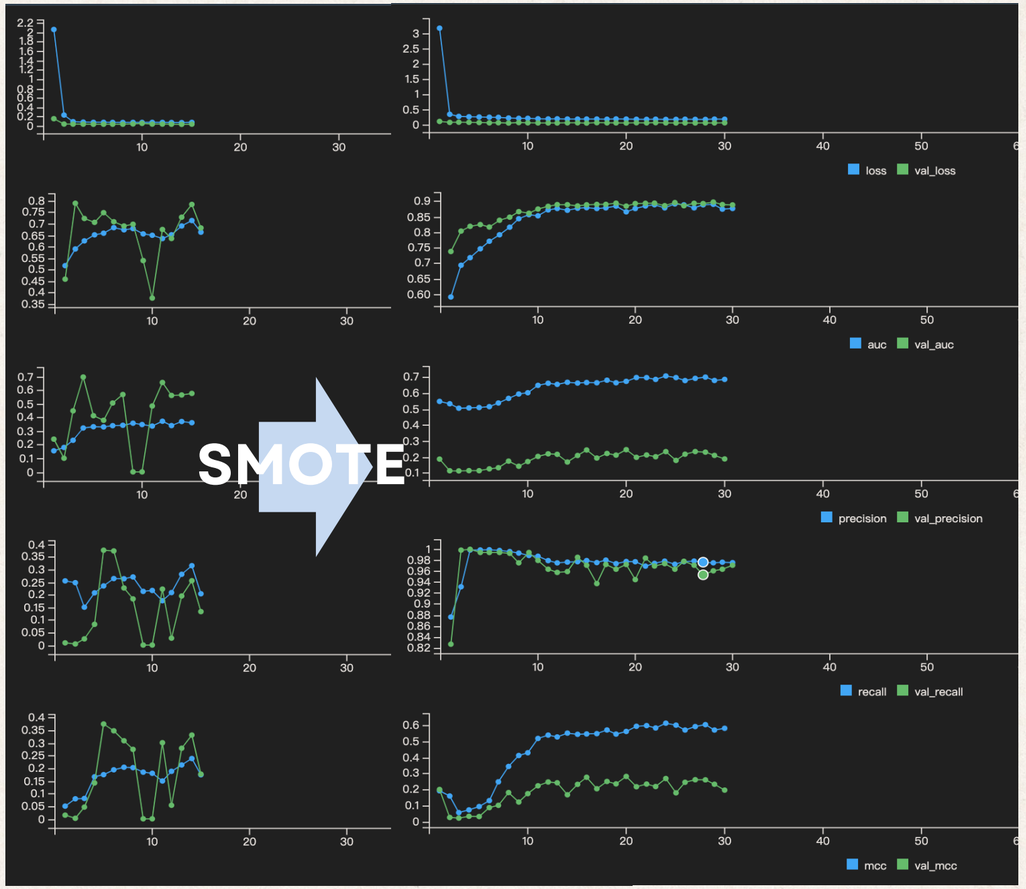
- **Test MCC**: 0.319
- **Test AUC**: 0.748

SMOTE →

- **Test MCC**: 0.473
- **Test AUC**: 0.887

Insights

- *SMOTE Oversampling:*
  - 1. Enhance class balance handling & discriminative ability
  - 2. Training curves show more stable validation metrics, reflecting stronger generalization & reduced bias toward majority class



Model more balanced & without severe overfitting

Neural Network Model

5-Fold Cross Validation(Results)

- **Mean MCC**: 0.502
- **Mean AUC**: 0.904

Fold	MCC	AUC
1	0.479	0.902
2	0.512	0.909
3	0.517	0.903
4	0.505	0.905
5	0.497	0.901

Insights

- ***1. MCC:*** a favorable value under imbalanced data, indicating relatively accurate predictions for both positive & negative samples

- ***2. Model Stability:*** AUC & MCC fluctuates within a very small range, stable model performance

Neural Network Model

Error Type(False Negative/Positive) Analysis

- **False Negative:** potential customers whom model predicts will not subscribe
- **False Positive:** customers predicted to subscribe but do not actually subscribe

		PREDICTED CLASS	
		Class=Yes	Class=No
ACTUAL CLASS	Class=Yes	a (TP)	b (FN)
	Class=No	c (FP)	d (TN)

Neural Network Model

Error Type(False Negative/Positive) Analysis(Insights)

- ***1. duration:*** Median & overall distribution of FP group are significantly higher
  - the higher the duration, the more likely the model is to predict “will subscribe”
  - while FN has shorter durations → the model overly relies on duration as predictive cue

- ***2. pdays:*** FP exhibits greater fluctuation & higher upper quantile
  - Customers with prolonged inactivity are prone to be misclassified as “will subscribe”
  - lead by the extreme values of pdays