Predicting Customer Churn Using Logistic Regression and Shallow ANN

By Pascal Opara

Agenda

3. Use Case Overview
4. Dataset Summary
5. Data Preparation & EDA
6. Model 1 – MV Logistic Regression
7. Model 2 – Shallow ANN
8. Visualization(Shallow ANN)
9. Conclusion
10. Q & A
11. Comparision Table

Use Case Overview

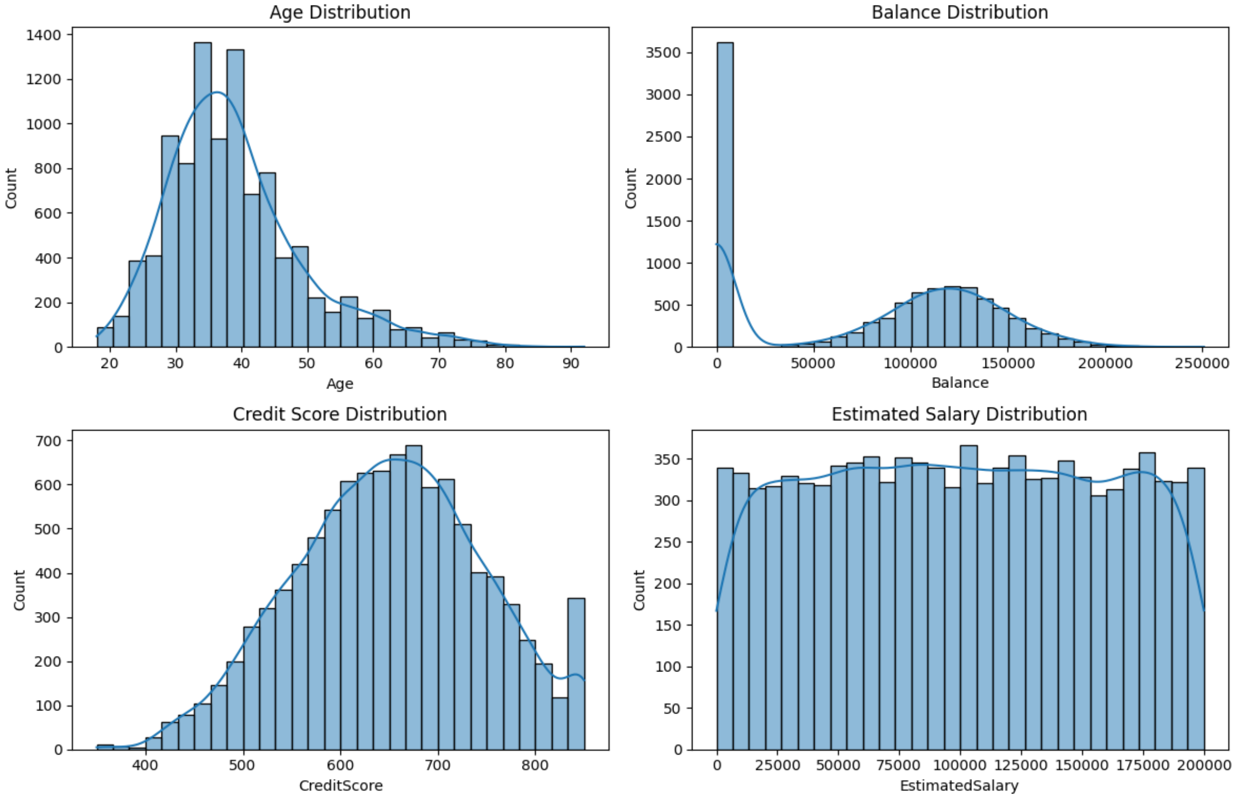
- Industry: Banking
- Problem: Customer churn leads to revenue loss
- Goal: Predict which customers are likely to churn
- Target Variable: Exited (0 = stayed, 1 = churned)

Dataset Summary

- Total records: 10000
- Features used:
  - Credit Score, Age, Tenure, Balance
  - Number of Products, Is ActiveMember
  - Geography, Gender(Encoded)
- Target: Exited

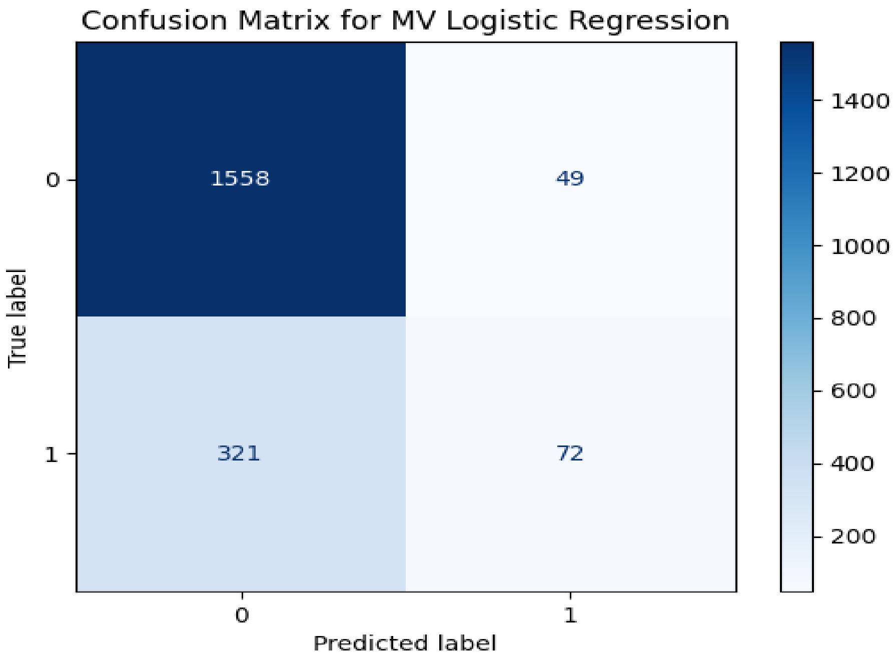
Data Preparation & EDA

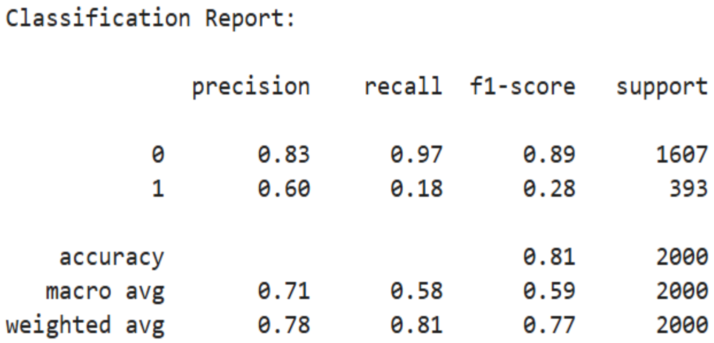
- Dropped irrelevant columns: RowNumber, CustomerId, Surname
- Label encoded Gender & Geography
- Scaled all X feautres using StandardScaler
- 80/20 train-test split



Model 1 – MV Logistic Regression

- Simple, interpretable model
- Output = probability of churn
- Used sigmoid + binary cross-entropy
- Achieved test accuracy: 81.55%
- Easy to interpret, fast to train.
- Imbalance issue: 1607:393



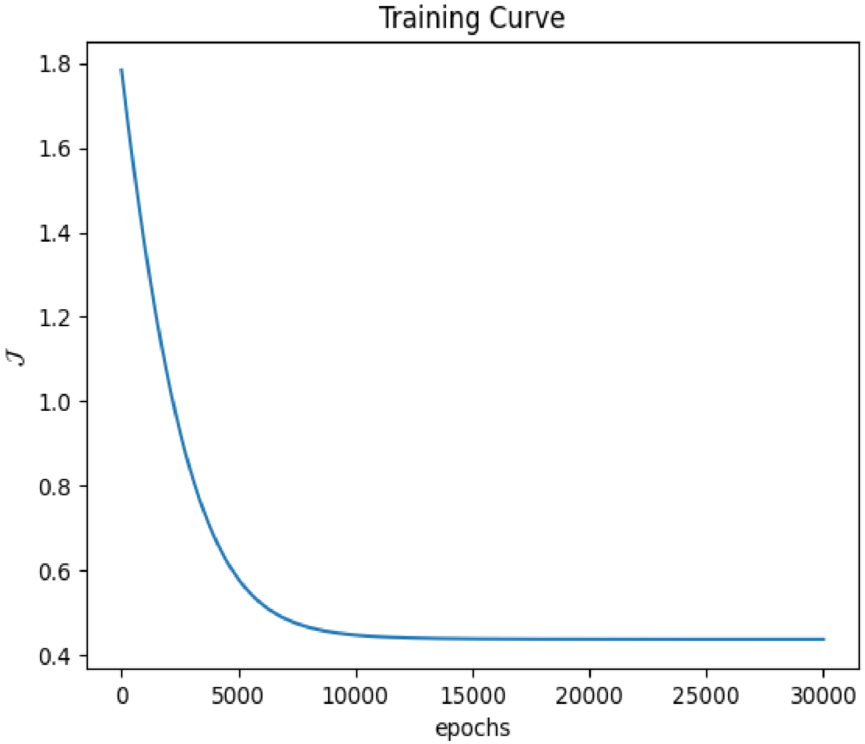


```
Classification Report:

              precision    recall  f1-score   support

           0       0.83      0.97      0.89      1607
           1       0.60      0.18      0.28       393

    accuracy                           0.81      2000
   macro avg       0.71      0.58      0.59      2000
weighted avg       0.78      0.81      0.77      2000
```



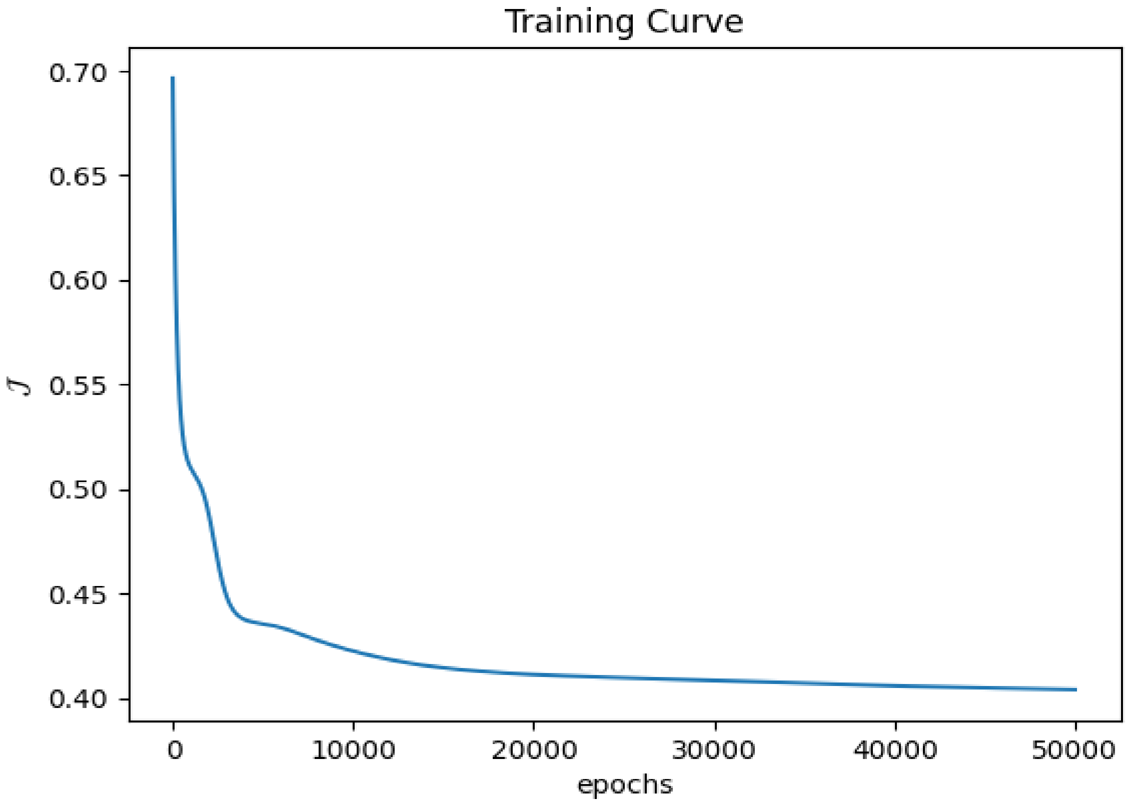
Model 2 – Shallow ANN

- Custom-built 2- layer network:

- Hidden Layer: tanh activation

- Output Layer: 1 neuron with sigmoid

- Loss: Binary cross-entropy
- Epochs: 5e4| Learning rate: 1e-2
- Weights initialized small to prevent saturation
- Loss Curve - Smooth convergence, stable training.
- Accuracy: 84.5%

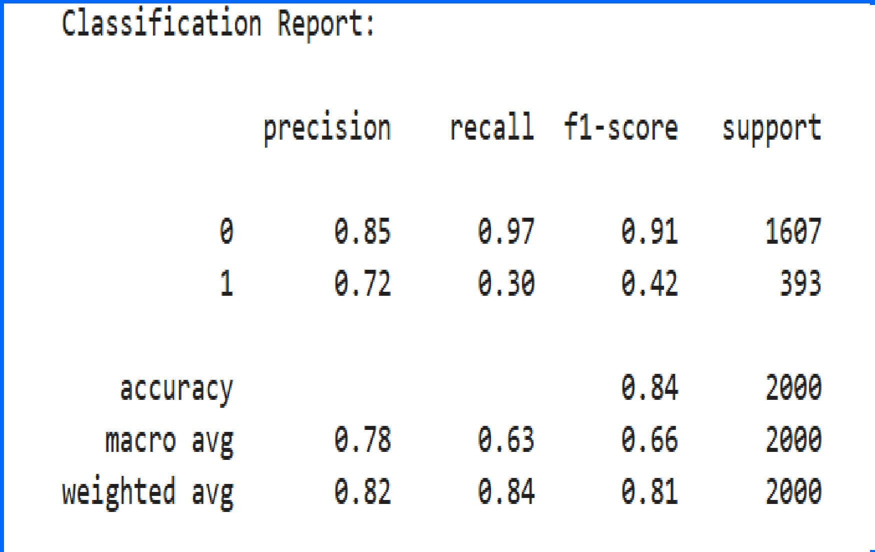


Visualization (Shallow ANN)

- 1561 – The model correctly predicted customers who did not churn
- 117 – The model correctly predicted customers who churned

- Precision (class 1) = 72% were correct
- Recall (class 1) = 30% was not detected.

Key Observation:
Strong performance on customers who did not churn than otherwise
Class imbalance – 393 out of 2000 churned

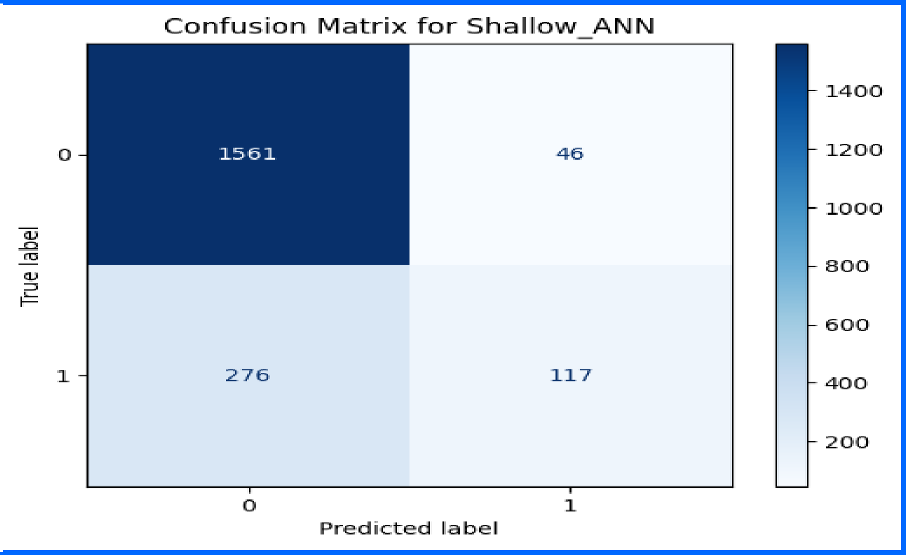


```
Classification Report:

              precision    recall  f1-score   support

           0       0.85      0.97      0.91      1607
           1       0.72      0.30      0.42       393

    accuracy                           0.84      2000
   macro avg       0.78      0.63      0.66      2000
weighted avg       0.82      0.84      0.81      2000
```



Conclusion

- Both models can predict churn
- MVLR is simple, ANN is flexible
- ANN gave better accuracy but might need more tuning

Future Steps:

- Use more neurons
- Try better weight initialization
- Evaluate class balance
- Solve imbalance issue

Thank you

Q & A

Comparision Table: Original vs Modified Shallow ANN

Original(Multiclass)	Modified(Binary)	Reason for Change
K = number of classes	K = 1	Binary classification needs only one output neuron
One_hot_encode(y)	y.reshape(-1, 1)	No need for one_hot encoding in binary classification
Softmax in output layer	Sigmoid in output layer	Sigmoid outputs a single probability between 0 and 1
Cross_entropy() loss	Binary_cross_entropy() loss	Binary cross-entropy is appropriate for 2-class problems
Argmax() for prediction	(Z[L] > 0.5).astype(int)	Matches outshape for binary classification