MLflow + Evidently AI: Student Lab – 65070507238

1. A brief description of the model, features, and metrics you tracked.

1.1 Model Description:

- **Type**: Logistic Regression (binary classification)
- **Architecture**: A scikit-learn Pipeline combining:
  - **StandardScaler**: Preprocesses features by standardizing them (zero mean, unit variance)
  - **LogisticRegression**: Binary classifier with default L2 regularization

1.2 Hyperparameters Tracked:

- model: "LogisticRegression" scaler: "StandardScaler"
- penalty: "l2" (L2 regularization) C: 1.0 (inverse regularization strength)
- max_iter: 1000

1.3 Features:

- The exact features are loaded from train.csv and test.csv (all columns except the target column)
- Features are standardized using StandardScaler before being fed to the classifier

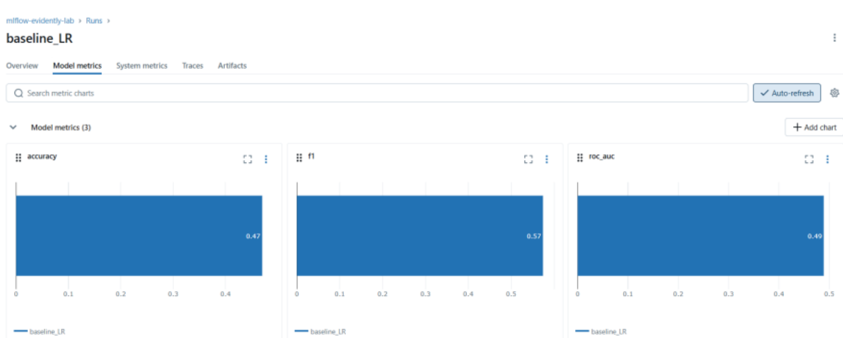
1.4 Metrics Tracked:

- **accuracy_score**: Measures the proportion of correct predictions
- **f1_score**: Harmonic mean of precision and recall (useful for imbalanced datasets)
- **roc_auc_score**: Area under the ROC curve (evaluates the model's ability to distinguish between classes at different probability thresholds)

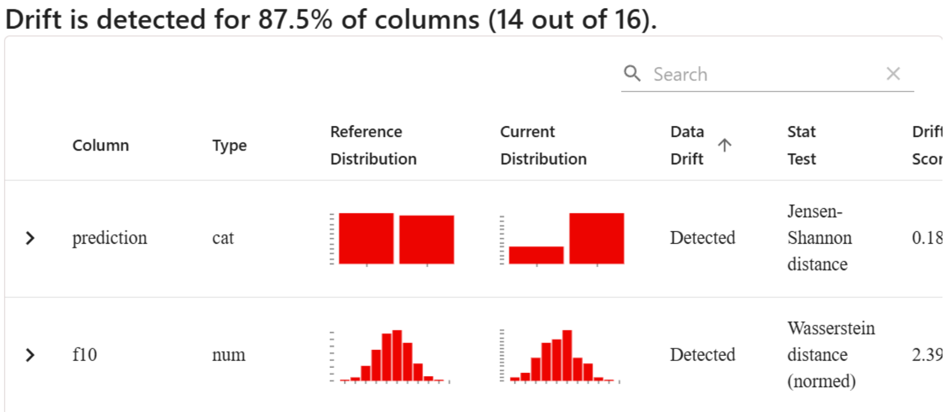
1.5 Visualizations Generated:

1. **Confusion Matrix Plot**: Shows true positives, true negatives, false positives, and false negatives
2. **ROC Curve**: Plots true positive rate vs. false positive rate

2. 1–2 screenshots from MLflow showing your metrics/plots.



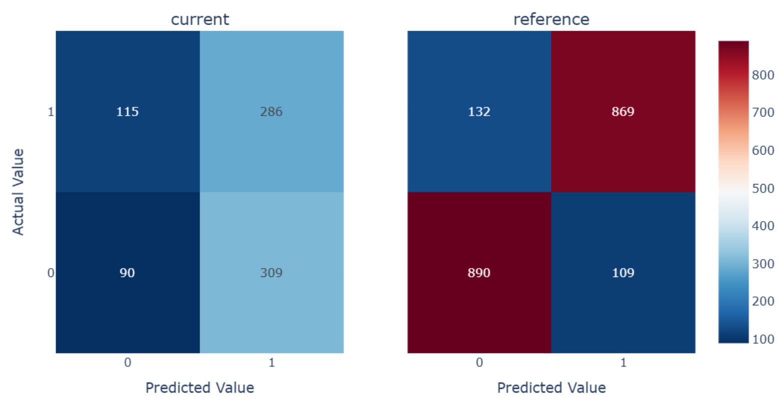
3. 1 screenshot from the Evidently HTML report highlighting any **data** or **target** drift.



Drift is detected for 87.5% of columns (14 out of 16).

	Column	Type	Reference Distribution	Current Distribution	Data Drift	Stat Test	Drift Scor
>	prediction	cat			Detected	Jensen-Shannon distance	0.18
>	f10	num			Detected	Wasserstein distance (normed)	2.39

4. Your interpretation: What changed between reference and current data? How might that impact model performance? What would you do next in production?



The change between reference and current data is

- **Data Drift**: Production data distribution differs from training
- **Concept Drift**: The relationship between features and target has changed
- **Seasonality/Temporal Effects**: Time-dependent patterns in new data
- **Class Imbalance**: Model may need retraining with updated class weights

That change might cause the impact model like

- Lower recall (missing important positive cases)
- Model may be less useful for critical applications requiring high sensitivity
- Metrics trained in reference data no longer representative of current performance

Next steps in production are to generate monitoring alerts for metric degradation and investigate root cause analysis on why class distribution shifted. Also, check if business requirements still align with current model performance

Sawit Koseeyaumporn 65070507238