1. Accuracy

Mengukur proporsi prediksi yang benar terhadap keseluruhan data.

Rumus:

* TP (True Positive): Prediksi churn yang benar.
* TN (True Negative): Prediksi tidak churn yang benar.
* FP (False Positive): Prediksi churn tetapi sebenarnya tidak churn.
* FN (False Negative): Prediksi tidak churn tetapi sebenarnya churn.

Kelemahan: Accuracy bisa menyesatkan jika data tidak seimbang (misalnya, jika mayoritas pelanggan tidak churn).

1. Precision

Mengukur akurasi prediksi churn dari total prediksi churn.

Rumus:

* Precision tinggi menunjukkan model jarang memberikan prediksi positif yang salah.

1. Recall (Sensitivity/True Positive Rate)

Mengukur kemampuan model dalam mendeteksi churn yang sebenarnya terjadi.

Rumus:

* Recall tinggi berarti model dapat menangkap sebagian besar kasus churn yang sebenarnya.

1. F1-Score

Merupakan rata-rata harmonis dari Precision dan Recall, memberikan keseimbangan antara keduanya.

Rumus:

* F1-Score lebih relevan jika data tidak seimbang.

1. F2-Score

Varian dari F1-Score yang memberikan bobot lebih besar pada Recall dibandingkan Precision.  
Rumus:

* Untuk F2-Score, β = 2, sehingga Recall menjadi 2 kali lebih penting dibanding Precision.

1. AUC-ROC (Area Under the Curve - Receiver Operating Characteristic)

Mengukur kemampuan model membedakan antara kelas churn dan tidak churn berdasarkan probabilitas prediksi.

* ROC Curve: Grafik antara True Positive Rate (TPR) dan False Positive Rate (FPR).
* AUC (Area Under the Curve): Nilai numerik yang menunjukkan luas di bawah kurva ROC. Nilai AUC mendekati 1 menunjukkan model yang sangat baik.

1. Log Loss (Logarithmic Loss)

Mengukur seberapa jauh prediksi probabilitas dari label yang benar, dengan memberikan penalti lebih tinggi untuk prediksi yang salah dengan probabilitas tinggi.  
Rumus:

* N: Jumlah sampel.
* yi: Label aktual (1 untuk churn, 0 untuk tidak churn).
* pi: Probabilitas prediksi untuk kelas 1 (churn).

Log Loss lebih sensitif terhadap probabilitas yang salah dibandingkan metrik lain seperti Accuracy, sehingga sangat berguna untuk model probabilistik.

* Accuracy: Bagus untuk data seimbang.
* Precision & Recall: Penting jika salah satu kesalahan (False Positive atau False Negative) lebih kritis.
* F1-Score & F2-Score: Cocok untuk dataset tidak seimbang.
* AUC-ROC: Mengukur kemampuan pemisahan kelas.
* Log Loss: Digunakan untuk model probabilistik.