"Studi Komparatif Gradient Boosting dan CatBoost dalam Memprediksi Churn di Apartemen Trillium"

1. Memahami Sumber Masalah Penelitian

Sumber masalah dalam penelitian ini yaitu pelanggan tidak melanjutkan sewa atau berhenti kontrak (churn) yang dapat menurunkan pendapatan dan meningkatkan biaya operasional untuk menarik pelanggan baru. Oleh karena itu, memprediksi churn pelanggan secara akurat dapat membantu manajemen dalam mengambil keputusan yang lebih baik terkait strategi pemasaran dan retensi pelanggan. Penelitian ini bertujuan untuk membandingkan kinerja dua algoritma machine learning, yaitu Gradient Boosting dan CatBoost, dalam memprediksi churn pelanggan berdasarkan data historis.

1. Merumuskan Masalah Penelitian

* Bagaimana kinerja algoritma Gradient Boosting dan CatBoost untuk memprediksi churn di Apartemen Trillium?
* Dari kedua algoritma tersebut, manakah yang paling baik untuk mendukung retensi pelanggan?

1. Membuat Kerangka Konsep

Kerangka konsep penelitian ini dapat digambarkan melalui beberapa variabel yang saling berhubungan:

* Input Variabel (Fitur): Data historis pelanggan, seperti durasi sewa, jumlah pembayaran, frekuensi keterlambatan pembayaran, umur pelanggan, lokasi unit, dan interaksi dengan fasilitas.
* Proses Prediksi: Penggunaan dua algoritma, yaitu Gradient Boosting dan CatBoost, untuk memprediksi kemungkinan pelanggan akan churn.
* Output: Hasil prediksi churn (ya/tidak) serta evaluasi kinerja dari kedua algoritma berdasarkan metrik seperti akurasi, precision, recall, dan AUC.

1. Merumuskan Hipotesis

Hipotesis dalam penelitian ini bisa dirumuskan sebagai berikut:

* Hipotesis Utama (H1): Algoritma CatBoost memiliki akurasi yang lebih tinggi daripada Gradient Boosting dalam memprediksi churn pelanggan di Apartemen Trillium.
* Hipotesis Nol (H0): Tidak ada perbedaan signifikan dalam akurasi antara algoritma Gradient Boosting dan CatBoost dalam memprediksi churn pelanggan di Apartemen Trillium.

1. Mendesain Penelitian
2. Studi Literatur: Meneliti penelitian sebelumnya mengenai churn prediction dan membandingkan algoritma Gradient Boosting dan CatBoost dalam konteks prediksi churn pada sektor apartemen.
3. Pengumpulan Data: Mengumpulkan data transaksi dan riwayat sewa dari Apartemen Trillium untuk dianalisis dalam model prediksi churn.
4. Pre-processing: Melakukan pembersihan, transformasi, dan normalisasi data serta membagi data menjadi data pelatihan dan pengujian.
5. Pemodelan: Menerapkan algoritma Gradient Boosting dan CatBoost pada data untuk memprediksi kemungkinan churn pelanggan.
6. Evaluasi: Mengukur kinerja model menggunakan metrik seperti akurasi, precision, recall, F1-score, dan AUC-ROC untuk membandingkan efektivitas kedua model.
7. Visualisasi: Menyajikan hasil evaluasi model melalui visualisasi seperti confusion matrix dan AUC-ROC curve untuk mempermudah interpretasi hasil.
8. Mengumpulkan & Menganalisis Data

Mengumpulkan Data: Data pelanggan dari Apartemen Trillium akan dikumpulkan, yang meliputi riwayat sewa, pembayaran, dan informasi lainnya yang relevan dengan churn.

1. Membuat Kesimpulan

Berdasarkan hasil analisis, kesimpulan dapat dibuat mengenai:

* Model Terbaik: Algoritma mana yang lebih unggul dalam memprediksi churn pelanggan di Apartemen Trillium, baik dari segi akurasi, efisiensi, maupun interpretabilitas.
* Faktor-Faktor Kritis: Menentukan faktor-faktor penting yang mempengaruhi prediksi churn, seperti keterlambatan pembayaran atau lama tinggal pelanggan.
* Rekomendasi Praktis: Memberikan rekomendasi kepada manajemen Apartemen Trillium terkait strategi untuk mengurangi churn berdasarkan temuan penelitian.