

# Analisis Siklus Hidup Pelanggan dan Prediksi Churn Menggunakan Model Rantai Markov Waktu Diskrit

## Kelompok 4

Muhammad Zaki Z

Ilmu Aktuaria

muhammad.zaki22@ui.ac.id

Joshua Novandri

Ilmu Aktuaria

joshua.novandri@ui.ac.id

Stephanie Febriana

Ilmu Aktuaria

stephaniefbriana@sci.ui.ac.id

Andi Rizqy Ariqah

Ilmu Aktuaria

andi.rizqy@ui.ac.id

Devina Vanessa H

Ilmu Aktuaria

devina.vanessa@ui.ac.id

## Abstrak

Retensi pelanggan merupakan tantangan krusial dalam industri telekomunikasi yang kompetitif. Model prediksi *churn* konvensional sering kali gagal menangkap dinamika perilaku pelanggan yang berubah seiring waktu. Penelitian ini mengusulkan dan mengimplementasikan model Rantai Markov Waktu Diskrit (DTMC) *multi-state* untuk memodelkan perjalanan atau siklus hidup pelanggan secara lebih granuler. Dengan menggunakan dataset publik *Telco Customer Churn*, pelanggan diklasifikasikan ke dalam empat status: 'New' (1 – 12 bulan), 'Established' (13 – 48 bulan), 'Loyal' (> 48 bulan), dan 'Churn' sebagai status penyerap. Dari data tersebut, matriks probabilitas transisi bulanan berhasil diestimasi. Hasil utama menunjukkan bahwa pelanggan pada fase 'New' memiliki risiko *churn* tertinggi secara signifikan ( $\approx 47.68\%$  per bulan), sementara pelanggan 'Loyal' menunjukkan inersia yang sangat tinggi untuk bertahan ( $\approx 86.72\%$  per bulan). Analisis jangka panjang menggunakan perpangkatan matriks memprediksi bahwa seorang pelanggan baru memiliki probabilitas kumulatif untuk *churn* yang dapat mencapai lebih dari ( $\approx 95.83\%$ ) dalam 6 bulan pertama dan meningkat hingga ( $\approx 99.96\%$ ) setelah tiga tahun. Model ini membuktikan bahwa pendekatan berbasis siklus hidup tidak hanya mampu mengidentifikasi segmen pelanggan paling rentan, tetapi juga memberikan wawasan kuantitatif untuk perancangan strategi retensi yang lebih efektif dan tertarget.

**Keywords:** *Churn* Pelanggan, Proses Stokastik, Rantai Markov Waktu Diskrit, Siklus Hidup Pelanggan, Telekomunikasi.

# 1 Pendahuluan

## 1.1 Latar Belakang Masalah

Industri telekomunikasi modern ditandai oleh tingkat persaingan yang sangat tinggi, di mana biaya untuk mengakuisisi pelanggan baru seringkali jauh lebih besar daripada biaya untuk mempertahankan pelanggan yang sudah ada. Oleh karena itu, retensi pelanggan menjadi faktor krusial bagi profitabilitas dan keberlanjutan bisnis perusahaan. Fenomena pelanggan yang berhenti berlangganan, atau dikenal sebagai *customer churn*, merupakan tantangan utama yang dihadapi industri ini. Memahami dinamika di balik keputusan pelanggan untuk *churn* adalah kunci untuk merancang strategi retensi yang efektif.

Perilaku pelanggan tidaklah homogen; seorang pelanggan yang baru berlangganan selama dua bulan memiliki karakteristik dan risiko churn yang berbeda secara signifikan dibandingkan pelanggan yang telah setia selama lima tahun. Namun, banyak model prediksi *churn* memperlakukan semua pelanggan secara seragam, sehingga kehilangan wawasan penting terkait siklus hidup (*lifecycle*) pelanggan. Proses stokastik, khususnya Rantai Markov Waktu Diskrit (DTMC), menyediakan kerangka kerja matematis yang kuat untuk memodelkan sistem yang bertransisi antar status dari waktu ke waktu dalam kondisi ketidakpastian (Ross, 2014). Berangkat dari (Pfeifer and Carraway, 2000) dengan mendefinisikan status berdasarkan fase *lifecycle* pelanggan, DTMC memungkinkan kita untuk memetakan dan menganalisis perjalanan pelanggan secara lebih granuler (Grimmett and Stirzaker, 2001) dan dinamis (Gusev, 2019).

## 1.2 Rumusan Masalah

Penelitian ini berfokus pada penerapan model DTMC untuk menganalisis *lifecycle* pelanggan telekomunikasi. Berdasarkan latar belakang tersebut, rumusan masalah yang hendak dijawab adalah sebagai berikut:

- Bagaimana cara memodelkan perjalanan pelanggan telekomunikasi melalui berbagai fase loyalitas (*New, Established, Loyal*) hingga berhenti berlangganan (*Churn*) menggunakan Rantai Markov Waktu Diskrit?
- Berapakah estimasi probabilitas transisi bulanan bagi seorang pelanggan untuk berpindah antar fase loyalitas atau menuju status *Churn*?
- Bagaimana perbedaan risiko *churn* pada setiap fase *lifecycle* pelanggan, dan fase manakah yang memiliki risiko tertinggi?
- Berapakah ekspektasi rata-rata masa berlangganan (*customer lifetime*) seorang pelanggan berdasarkan fase loyalitas mulanya?

## 1.3 Tujuan Penelitian

Sejalan dengan rumusan masalah, tujuan dari penelitian ini adalah:

- Mengembangkan model Rantai Markov Waktu Diskrit (DTMC) *multi-state* yang merepresentasikan siklus hidup pelanggan telekomunikasi.
- Mengestimasi matriks probabilitas transisi bulanan dari model tersebut menggunakan dataset publik *Telco Customer Churn*.

- Menganalisis dinamika *churn* dengan mengidentifikasi fase-fase paling kritis dalam perjalanan pelanggan.
- Menghitung metrik bisnis strategis, seperti probabilitas retensi jangka panjang dan ekspektasi masa berlangganan pelanggan berdasarkan fase loyalitas mulanya, sebagai hasil dari analisis model.

## 1.4 Manfaat Penelitian

Hasil dari penelitian ini diharapkan dapat memberikan manfaat sebagai berikut:

- **Manfaat Praktis/Industri:** Menyediakan sebuah model kuantitatif bagi perusahaan telekomunikasi untuk memahami perilaku pelanggan secara lebih mendalam. Wawasan yang dihasilkan dari model ini dapat menjadi dasar untuk merancang strategi retensi yang lebih tertarget dan efisien, misalnya dengan memfokuskan sumber daya pada pelanggan di fase yang paling berisiko terhadap *churn*.
- **Manfaat Akademis/Teoritis:** Menjadi studi kasus nyata penerapan teori proses stokastik, khususnya DTMC, pada permasalahan bisnis. Penelitian ini menunjukkan bagaimana konsep-konsep teoretis seperti matriks transisi dan state penyerap dapat diimplementasikan untuk menghasilkan wawasan yang dapat ditindaklanjuti secara praktis.

## 1.5 Batasan Masalah

Untuk menjaga fokus penelitian, terdapat beberapa batasan masalah yang ditetapkan:

- Penelitian ini menggunakan dataset *snapshot* (cross-sectional), sehingga transisi antar status bersifat inferensi (disimpulkan), bukan observasi langsung dari data panel.
- Probabilitas transisi diasumsikan konstan untuk semua pelanggan di dalam satu grup *lifecycle* yang sama.
- Model tidak memperhitungkan faktor eksternal lain yang dapat mempengaruhi *churn*, seperti kampanye pemasaran, perubahan harga dari kompetitor, atau kondisi ekonomi makro.
- Batasan fase *lifecycle* (1-12 bulan, 13-48 bulan, > 48 bulan) didefinisikan untuk keperluan studi ini dan dapat disesuaikan pada penelitian selanjutnya.

Fenomena stokastik muncul dalam berbagai bidang, seperti antrian layanan, telekomunikasi, dan sistem biologis (Ross, 2014). Studi ini bertujuan untuk mengimplementasikan model stokastik yang sesuai dalam memodelkan suatu masalah riil, serta menganalisis karakteristik dan hasil simulasi model tersebut.

## 2 Metodologi

Metodologi penelitian ini disusun untuk memastikan implementasi model stokastik yang sistematis dan dapat direproduksi. Tahapan metodologi mencakup deskripsi dan pemilihan model, definisi ruang keadaan dan parameter, formulasi matematis, serta metode analisis data.

## 2.1 Deskripsi Model

Model stokastik yang digunakan dalam penelitian ini adalah **Rantai Markov Waktu Diskrit (Discrete-Time Markov Chain, DTMC)**. Model ini dipilih karena kemampuannya yang sangat baik dalam merepresentasikan sistem yang bertransisi di antara serangkaian status yang berbeda pada interval waktu yang diskrit (yaitu, dari bulan ke bulan). Karakteristik utama yang mendasari DTMC adalah **Sifat Markov (Markov Property)**, yang menyatakan bahwa probabilitas transisi ke status berikutnya hanya bergantung pada status saat ini, dan tidak pada urutan status yang mendahuluinya. Properti ini sangat sesuai untuk memodelkan *churn* pelanggan, di mana keputusan pelanggan di masa depan dapat diasumsikan sangat dipengaruhi oleh fase loyalitas mereka saat ini.

## 2.2 Pengumpulan Data

Penelitian ini menggunakan data sekunder yang bersumber dari dataset publik “**Telco Customer Churn**” yang tersedia di platform *Kaggle* (Blastchar, 2018). Dataset ini berisi 7.043 data pelanggan dengan 21 atribut, yang mencakup informasi demografis, layanan yang digunakan (telepon, internet, TV streaming, dll.), detail kontrak, serta status *churn* pelanggan. Atribut kunci yang menjadi dasar pembentukan status dalam model ini adalah ‘tenure’, yang merepresentasikan jumlah bulan pelanggan telah berlangganan.

## 2.3 Definisi Ruang Keadaan dan Parameter

Untuk memodelkan perjalanan pelanggan, ruang keadaan (*state space*)  $\mathcal{S}$  didefinisikan berdasarkan fase siklus hidup pelanggan. Model ini memiliki empat status:

1. **State 1: New (Baru):** Pelanggan dengan masa berlangganan (*tenure*) 1–12 bulan.
2. **State 2: Established (tetap):** Pelanggan dengan *tenure* 13–48 bulan.
3. **State 3: Loyal (Setia):** Pelanggan dengan *tenure* lebih dari 48 bulan.
4. **State 4: Churn (Berhenti):** Status penyerap (*absorbing state*) bagi pelanggan yang telah berhenti berlangganan.

Parameter utama model adalah probabilitas transisi bulanan  $p_{ij}$ , yang menyatakan probabilitas berpindah dari status  $i$  ke status  $j$  dalam satu bulan. Asumsi yang digunakan adalah probabilitas transisi ini stasioner (tidak berubah seiring waktu) untuk setiap kelompok status.

## 2.4 Formulasi Matematis

Probabilitas transisi satu langkah dari status  $i$  ke status  $j$  didefinisikan sebagai:

$$p_{ij} = P(X_{n+1} = j | X_n = i)$$

di mana  $X_n$  adalah status sistem pada waktu (bulan) ke- $n$ . Seluruh probabilitas ini dirangkum dalam sebuah Matriks Probabilitas Transisi ( $P$ ) berukuran  $4 \times 4$ :

$$P = \begin{pmatrix} p_{11} & p_{12} & p_{13} & p_{14} \\ p_{21} & p_{22} & p_{23} & p_{24} \\ p_{31} & p_{32} & p_{33} & p_{34} \\ p_{41} & p_{42} & p_{43} & p_{44} \end{pmatrix} = \begin{pmatrix} p_{\text{New} \rightarrow \text{New}} & p_{\text{New} \rightarrow \text{Est}} & p_{\text{New} \rightarrow \text{Loyal}} & p_{\text{New} \rightarrow \text{Churn}} \\ p_{\text{Est} \rightarrow \text{New}} & p_{\text{Est} \rightarrow \text{Est}} & p_{\text{Est} \rightarrow \text{Loyal}} & p_{\text{Est} \rightarrow \text{Churn}} \\ p_{\text{Loyal} \rightarrow \text{New}} & p_{\text{Loyal} \rightarrow \text{Est}} & p_{\text{Loyal} \rightarrow \text{Loyal}} & p_{\text{Loyal} \rightarrow \text{Churn}} \\ p_{\text{Churn} \rightarrow \text{New}} & p_{\text{Churn} \rightarrow \text{Est}} & p_{\text{Churn} \rightarrow \text{Loyal}} & p_{\text{Churn} \rightarrow \text{Churn}} \end{pmatrix}$$

Elemen matriks  $p_{ij}$  diestimasi dari data. Misalkan  $c_i$  adalah rasio *churn* empiris pada grup status  $i$  dan  $d_i$  adalah durasi (dalam bulan) dari status transient  $i$ . Maka:

- Probabilitas *Churn*:  $p_{i,4} = c_i$
- Probabilitas Bertahan dan Lulus ke Status Berikutnya:  $p_{i,i+1} = \frac{1}{d_i} \times (1 - c_i)$
- Probabilitas Bertahan dan Tetap di Status yang Sama:  $p_{ii} = \frac{d_i-1}{d_i} \times (1 - c_i)$
- Untuk state penyerap *Churn* (status 4):  $p_{44} = 1$  dan  $p_{4j} = 0$  untuk  $j \neq 4$ .
- Probabilitas transisi lainnya yang tidak mungkin (misal, dari ‘Established’ ke ‘New’) bernilai 0.

Distribusi probabilitas status setelah  $n$  langkah (bulan) dapat dihitung dengan persamaan:

$$\pi^{(n)} = \pi^{(0)} P^n$$

di mana  $\pi^{(n)}$  adalah vektor baris distribusi status pada waktu ke- $n$  dan  $\pi^{(0)}$  adalah distribusi status awal.

## 2.5 Metode Analisis

Analisis data dan implementasi model dilakukan menggunakan perangkat lunak **Python** versi 3.x dengan memanfaatkan pustaka (*libraries*) utama sebagai berikut:

- **Pandas**: Digunakan untuk memanipulasi data, membersihkan, serta melakukan pra-pemrosesan seperti membuat kategori ‘tenure’ untuk mendefinisikan status pelanggan.
- **NumPy**: Digunakan untuk komputasi numerik, khususnya dalam membuat dan melakukan operasi aljabar linear pada matriks transisi, seperti perpangkatan matriks ( $P^n$ ).

Proses analisis meliputi estimasi parameter  $p_{ij}$  dari data, pembentukan matriks  $P$ , dan analisis perilaku jangka panjang dengan menghitung perpangkatan matriks untuk memprediksi distribusi pelanggan di masa depan.

## 3 Hasil dan Pembahasan

Pada bagian ini, disajikan hasil analisis dari model Rantai Markov Waktu Diskrit (DTMC) *lifecycle* pelanggan yang telah dikembangkan. Pembahasan difokuskan pada interpretasi matriks transisi, analisis perilaku jangka panjang, serta diskusi mengenai kekuatan dan keterbatasan model yang digunakan.

### 3.1 Matriks Transisi dan Risiko Churn per Fase

Estimasi parameter dari dataset menghasilkan Matriks Probabilitas Transisi ( $P$ ) bulanan sebagai berikut, yang merangkum seluruh dinamika perpindahan pelanggan antar status.

Table 1: Matriks Probabilitas Transisi Bulanan (P)

Dari	New	Established	Loyal	Churn
New	0.4796	0.0436	0.000	0.4768
Established	0.000	0.7423	0.0212	0.2364
Loyal	0.000	0.000	0.8672	0.0951
Churn	0.000	0.000	0.000	1.000

### 3.1.1 Interpretasi Hasil

Matriks pada Tabel 1 memberikan wawasan kuantitatif mengenai perilaku pelanggan di setiap fase:

- **Risiko Churn Tertinggi pada Fase Awal:** Probabilitas seorang pelanggan ‘New’ untuk *churn* dalam satu bulan ( $p_{14}$ ) adalah sebesar 47.68%. Angka ini secara signifikan lebih tinggi dibandingkan pelanggan ‘Established’ (23.64%) dan ‘Loyal’ (9.51%). Hal ini secara kuantitatif memvalidasi hipotesis bahwa pelanggan baru adalah segmen yang paling rentan dan memerlukan perhatian retensi khusus.
- **Inersia Pelanggan Setia:** Pelanggan pada fase ‘Loyal’ memiliki probabilitas yang tinggi untuk tetap loyal ( $p_{33} = 86.72\%$ ). Hal ini menunjukkan tingkat inersia yang besar dan mencerminkan keberhasilan perusahaan dalam mempertahankan segmen ini.
- **Transisi Kelulusan:** Probabilitas seorang pelanggan ‘New’ untuk bertahan dan ‘lulus’ atau berubah status menjadi ‘Established’ ( $p_{12}$ ) adalah sebesar 4.36%. Angka ini menandakan bahwa sebagian besar pelanggan baru tidak bertahan dalam waktu yang lama untuk berpindah ke fase berikutnya. Hal ini dapat menjadi indikator bahwa efektivitas proses *onboarding* pelanggan perlu ditingkatkan.
- **Sifat State Penyerap:** Baris terakhir matriks menunjukkan bahwa status ‘Churn’ merupakan state penyerap. Artinya, setelah seorang pelanggan berhenti berlangganan, mereka tidak akan kembali ke status aktif. Hal ini sejalan dengan teori rantai Markov, dimana semua pelanggan pada akhirnya akan terserap ke status *churn* ( $p_{44} = 1$ ).

## 3.2 Analisis Perilaku Jangka Panjang

Untuk memahami evolusi kohort pelanggan dari waktu ke waktu, dilakukan perhitungan perpangkatan matriks transisi ( $P^n$ ). Tabel 2 menunjukkan distribusi probabilitas seorang pelanggan yang awalnya ‘New’ akan berada di setiap status setelah 6, 12, 18, 24, 30, dan 36 bulan.

### 3.2.1 Interpretasi dan Perbandingan dengan Teori

Hasil simulasi jangka panjang ini sepenuhnya konsisten dengan teori rantai Markov yang memiliki state penyerap (Norris, 1998).

- **Probabilitas Churn Kumulatif:** Dapat dilihat bahwa probabilitas kumulatif seorang pelanggan baru untuk telah churn meningkat seiring waktu, dari 95.8% setelah 6 bulan menjadi 99.9% setelah 3 tahun. Angka ini memberikan gambaran realistis mengenai tingkat retensi jangka panjang.

Table 2: Distribusi Probabilitas Status Pelanggan Baru Setelah  $n$  Bulan

n Bulan	P(di New)	P(di Est.)	P(di Loyal)	P(di Churn)
6	0.0122	0.0258	0.0038	0.9583
12	0.0001	0.0046	0.0034	0.9918
18	0.0000	0.0008	0.0021	0.9971
24	0.0000	0.0001	0.0012	0.9987
30	0.0000	0.0000	0.0007	0.9993
36	0.0000	0.0000	0.0004	0.9996

- **Konvergensi ke State Penyerap:** Seiring dengan  $n \rightarrow \infty$ , probabilitas pelanggan berada di state transient ('New', 'Established', 'Loyal') akan menuju nol, dan probabilitas berada di state 'Churn' akan menuju 1. Ini adalah perilaku fundamental dari rantai Markov ireduksibel dengan state penyerap, di mana semua individu pada akhirnya akan terserap.
- **Puncak Loyalitas:** Model ini memungkinkan kita melihat perjalanan pelanggan. Seorang pelanggan baru mencapai puncak probabilitas untuk berada di fase 'Established' setelah 24 bulan, dan setelah itu probabilitas untuk berada di fase 'Loyal' mulai meningkat secara signifikan.

### 3.3 Kekuatan dan Keterbatasan Model

#### 3.3.1 Kekuatan

- **Granularitas Analisis:** Berbeda dengan model klasifikasi biner yang hanya menjawab "apakah" seorang pelanggan akan *churn*, model DTMC ini memberikan pemahaman yang lebih dalam mengenai "kapan" *churn* paling mungkin terjadi. Dengan membagi perjalanan pelanggan ke dalam fase 'New', 'Established', dan 'Loyal', model ini memudahkan identifikasi fase-fase di mana risiko *churn* paling tinggi, sehingga memungkinkan strategi intervensi yang lebih tertarget dan presisi berdasarkan fase siklus hidup pelanggan.
- **Interpretasi Intuitif:** Salah satu kekuatan terbesar dari Rantai Markov adalah kemudahan interpretasinya. Setiap elemen  $p_{ij}$  dalam matriks transisi memiliki makna bisnis yang langsung dapat dipahami sehingga jauh lebih mudah dikomunikasikan kepada para pemangku kepentingan, seperti manajer pemasaran atau direksi. Oleh karena itu, model ini membantu menghubungkan tim analisis data dengan pengambil keputusan secara lebih efektif.
- **Kuantifikasi Metrik:** Model DTMC ini berfungsi sebagai fondasi kuantitatif untuk menurunkan metrik-metrik bisnis yang lebih canggih. Secara spesifik, model ini menyediakan dua input yang krusial untuk perhitungan *Customer Lifetime Value* (CLV) lebih lanjut yaitu probabilitas retensi pada setiap periode dan ekspektasi sisa masa hidup pelanggan dari setiap fase. Dengan informasi ini, perusahaan bisa menilai bahwa pelanggan yang 'Loyal' bukan hanya bernilai tinggi karena pembelian mereka sekarang, tapi juga karena peluang mereka untuk terus bertahan ke depannya. Selain itu, teknik seperti Matriks Fundamental memungkinkan kita melihat rata-rata waktu yang dihabiskan pelanggan di tiap fase sebelum *churn*.

### 3.3.2 Keterbatasan

- **Asumsi Stasioneritas:** Model ini beroperasi di bawah asumsi fundamental bahwa probabilitas transisi antar status bersifat stasioner atau tidak berubah seiring berjalannya waktu. Asumsi ini membuat model jadi lebih sederhana, tapi mungkin kurang tepat untuk menggambarkan kondisi nyata di industri telekomunikasi yang cepat berubah. Misalnya, adanya promosi besar dari pesaing, teknologi baru seperti 5G, atau perubahan aturan pemerintah bisa mengubah cara pelanggan berperilaku, sehingga peluang transisinya juga berubah. Karena itu, model ini paling cocok digunakan saat kondisi pasar relatif stabil, namun akurasi bisa menurun jika terjadi perubahan besar di luar dugaan.
- **Inferensi dari Data Cross-Sectional:** Penelitian ini mengestimasi probabilitas transisi dari data *snapshot* (cross-sectional) dengan menggunakan ‘tenure’ sebagai proksi waktu. Pendekatan ini secara implisit mengasumsikan bahwa perilaku pelanggan yang saat ini memiliki *tenure*  $t$  bulan adalah representasi yang baik untuk perilaku pelanggan dengan *tenure*  $t - 1$  bulan di periode berikutnya. Asumsi ”kohort sintetis” ini mengabaikan kemungkinan adanya ”efek kohort”, di mana kelompok pelanggan yang mendaftar pada waktu yang berbeda (misalnya, saat ada promosi besar) mungkin memiliki karakteristik loyalitas yang berbeda. Untuk mendapatkan hasil yang lebih akurat, idealnya digunakan data panel (longitudinal) yang melacak pelanggan yang sama dari waktu ke waktu, sehingga transisi antar status bisa diamati secara langsung.
- **Homogenitas dalam Grup:** Dengan mengelompokkan pelanggan hanya berdasarkan ‘tenure’, model ini secara tidak langsung mengasumsikan bahwa semua pelanggan dalam satu status memiliki karakteristik yang sama. Padahal, kenyataannya ada banyak perbedaan di dalam tiap kelompok. Sebagai contoh, seorang pelanggan ‘New’ dengan kontrak dua tahun dan layanan premium (Fiber Optik, TV Streaming) memiliki profil risiko yang sangat berbeda dari pelanggan ‘New’ lain dengan layanan biasa dan kontrak bulanan. Model ini ”merata-ratakan” perilaku dari sub-segmen yang berbeda menjadi satu probabilitas transisi tunggal untuk setiap grup. Akibatnya, meskipun akurat secara agregat, model ini mungkin kurang presisi untuk memprediksi perilaku segmen pelanggan yang lebih spesifik.

## 4 Kesimpulan

Penelitian ini membuktikan bahwa model Rantai Markov Waktu Diskrit (DTMC) dapat diadaptasi secara efektif untuk menganalisis tantangan bisnis nyata yang lebih kompleks seperti retensi pelanggan. Dengan membagi perjalanan pelanggan telekomunikasi ke dalam fase-fase siklus hidup, pendekatan ini memungkinkan analisis yang lebih mendalam, bukan hanya memprediksi kemungkinan pelanggan akan *churn*, tetapi juga memperkirakan kapan dan dalam kondisi apa hal tersebut paling mungkin terjadi.

### Ringkasan Temuan Utama

Dari analisis yang telah dilakukan, beberapa temuan kunci berhasil diidentifikasi. Pertama, model DTMC 4-state secara efektif mengkuantifikasi bahwa risiko churn tidaklah seragam. Terbukti secara signifikan bahwa **pelanggan pada fase ‘New’ (1-12 bulan)**



adalah yang paling rentan, dengan probabilitas churn bulanan yang jauh melampaui pelanggan yang lebih tetap. Kedua, model ini berhasil memproyeksikan nasib sebuah kohort pelanggan baru dari waktu ke waktu, menunjukkan bahwa sebagian besar churn terjadi pada tahun-tahun awal, yang menggarisbawahi pentingnya kesan pertama dan pengalaman awal pelanggan. Temuan-temuan ini secara kolektif menegaskan bahwa pendekatan *lifecycle* adalah kunci untuk membuka pemahaman yang lebih otentik tentang perilaku pelanggan.

## Implikasi Praktis dan Teoritis

Wawasan yang dihasilkan dari model ini membawa implikasi yang dapat ditindaklanjuti. Dari sisi praktis, perusahaan telekomunikasi dapat merancang **strategi retensi yang lebih cerdas dan tertarget**. Alih-alih menyebar diskon secara acak, perusahaan dapat mengalokasikan sumber daya secara intensif pada 6-12 bulan pertama masa berlangganan untuk membangun loyalitas sejak dini. Model ini juga dapat menjadi alat bantu untuk peramalan bisnis, dengan memberikan proyeksi jumlah pelanggan aktif di masa depan yang lebih akurat. Secara teoritis, proyek ini menjadi sebuah studi kasus yang solid tentang bagaimana pemilihan dan pendefinisian ruang keadaan (*state space*) secara cermat dapat meningkatkan daya penjelasan sebuah model stokastik secara drastis.

## Saran dan Arah Penelitian Selanjutnya

Meskipun model ini memberikan hasil yang menjanjikan, masih terdapat ruang untuk pengembangan lebih lanjut. Beberapa saran yang dapat dipertimbangkan ke depannya antara lain:

- **Memperkaya Definisi Status:** Penambahan variabel lain seperti jenis kontrak, layanan yang digunakan, atau perilaku penggunaan dapat menghasilkan segmentasi yang lebih granular—misalnya membedakan antara pelanggan New (Fiber) dan New (DSL).
- **Menggunakan Model Stokastik Lanjutan:** Untuk menangkap dinamika yang lebih kompleks, penelitian di masa depan bisa mengeksplorasi Rantai Markov Waktu Kontinu (CTMC) untuk memodelkan *churn* yang bisa terjadi kapan saja, tidak hanya di akhir bulan.
- **Validasi dengan Data Panel:** Jika memungkinkan, menerapkan model ini pada data panel (longitudinal) akan memberikan validasi yang lebih kuat, karena memungkinkan pengamatan transisi pelanggan secara langsung alih-alih hanya mengandalkan estimasi dari data snapshot.

Penelitian ini menegaskan bahwa pendekatan kuantitatif berbasis DTMC dapat menjadi alat yang efektif dalam memahami pola retensi pelanggan secara sistematis. Dengan analisis yang tepat, model ini mampu menghasilkan wawasan yang relevan untuk mendukung pengambilan keputusan strategis berbasis data. Lampiran kode program dapat diakses melalui tautan berikut: <https://github.com/zakizulham/proyek-stokastik-churn-telekomunikasi>.

## References

- Blastchar (2018). Telco Customer Churn. Kaggle. Diakses pada 7 Juni 2025. Tersedia di: <https://www.kaggle.com/datasets/blastchar/telco-customer-churn>.
- Grimmett, G. R. and Stirzaker, D. R. (2001). *Probability and Random Processes*. Oxford University Press, 3rd edition.
- Gusev, M. (2019). A Markov chain-based model for customer churn prediction in telecommunication services. *Informatics*, 6(4):44.
- Norris, J. R. (1998). *Markov Chains*. Cambridge Series in Statistical and Probabilistic Mathematics. Cambridge University Press.
- Pfeifer, P. E. and Carraway, R. L. (2000). A simple model for predicting customer lifetime value using the Pareto/NBD and Markov chain models. *Journal of Interactive Marketing*, 14(2):47–58.
- Ross, S. M. (2014). *Introduction to Probability Models*. Academic Press, 11th edition.