

Usability Fitur dan Prediksi Retensi Pengguna TikTok Shop dengan Naïve Bayes

Feature-Level Usability and User Retention Prediction in TikTok Shop using Naïve Bayes

¹Jessica Gabriel Lantang, ²Evi Maria*

^{1,2}Program Studi Sistem Informasi, Fakultas Teknologi Informasi, Universitas Kristen Satya Wacana

^{1,2}Jl. Dr. O. Notohamidjodjo Blotongan, Sidorejo, Kota Salatiga, Jawa Tengah, Indonesia

*e-mail: evi.maria@uksw.edu

(received: 19 March 2026, revised: 1 April 2026, accepted: 4 April 2026)

Abstrak

Retensi pengguna merupakan faktor penting dalam keberlanjutan *platform social commerce*, namun hubungan antara *usability* dan retensi belum sepenuhnya konsisten dalam berbagai studi. Tujuan penelitian ini adalah untuk menganalisis pengaruh *usability* pada level fitur terhadap retensi pengguna TikTok Shop pada Generasi Z, serta mengidentifikasi pola retensi menggunakan pendekatan prediktif. Penelitian ini menggunakan pendekatan kuantitatif dengan desain eksplanatori-prediktif. Data dikumpulkan dari 115 responden melalui kuesioner berbasis *System Usability Scale* (SUS) untuk mengukur *usability* fitur *live shopping* dan *product reviews*. Analisis dilakukan menggunakan regresi logistik biner dan algoritma Naïve Bayes. Hasil penelitian menunjukkan bahwa kedua fitur berada pada kategori *usability marginal-low* dengan skor di bawah ambang kelayakan. Namun, hasil regresi logistik menunjukkan bahwa *usability* tidak berpengaruh signifikan terhadap retensi pengguna. Di sisi lain, model Naïve Bayes menunjukkan akurasi sebesar 88,64 persen dengan nilai *macro-average F1-score* sebesar 58,71 persen, namun memiliki keterbatasan dalam mendeteksi pengguna yang tidak bertahan, yang ditunjukkan nilai *recall* kelas minoritas sebesar 14,29 persen akibat ketidakseimbangan distribusi kelas. Ini artinya *usability* pada level fitur memiliki keterbatasan dalam menjelaskan retensi pengguna dalam konteks *social commerce*. Penelitian ini berkontribusi menegaskan bahwa retensi pengguna tidak sepenuhnya ditentukan oleh *usability*, serta menunjukkan bahwa pendekatan prediktif memberikan *insight* tambahan yang tidak terjabarkan dalam model eksplanatori.

Kata kunci: *usability* fitur, regresi logistik, naïve bayes, social commerce, retensi pengguna

Abstract

User retention is a critical factor in the sustainability of social commerce platforms; however, the relationship between usability and retention remains inconsistent across various studies. This study aims to analyze the effect of feature-level usability on user retention in TikTok Shop among Generation Z users, as well as to identify retention patterns using a predictive approach. The study adopts a quantitative approach with an explanatory–predictive research design. Data were collected from 115 respondents through a questionnaire based on the System Usability Scale (SUS), focusing on the usability of live shopping and product review features. The analysis was conducted using binary logistic regression and the Naïve Bayes algorithm. The results show that both features fall into the marginal-low usability category, with scores below the acceptable threshold. However, the logistic regression results indicate that usability does not have a significant effect on user retention. On the other hand, the Naïve Bayes model achieved an accuracy of 88.64%, with a macro-average F1-score of 58.71%, but showed limitations in detecting non-retained users, as indicated by a minority-class recall of 14.29% due to class imbalance. These findings suggest that feature-level usability has limited explanatory power in predicting user retention within the social commerce context. This study contributes by reinforcing the argument that user retention is not solely determined by usability and demonstrates that predictive approaches can provide additional insights that are not captured by explanatory models.

Keywords: *feature usability, logistic regression, naïve bayes, social commerce, user retention*

1 Pendahuluan

Retensi pengguna merupakan faktor kunci dalam keberlanjutan *platform* digital, khususnya dalam konteks *social commerce* yang menggabungkan interaksi sosial dan aktivitas transaksi dalam satu ekosistem. Retensi tidak hanya mencerminkan keberlanjutan penggunaan, tetapi juga menunjukkan kemampuan *platform* dalam mempertahankan nilai pengalaman pengguna secara konsisten dari waktu ke waktu [1]. Dalam perspektif sistem informasi, retensi pengguna (*continuance intention*), yaitu keputusan pengguna untuk terus menggunakan sistem berdasarkan evaluasi pengalaman pasca penggunaan [2]. Dengan demikian, keberlanjutan penggunaan tidak ditentukan oleh adopsi awal, tetapi oleh kualitas pengalaman yang dirasakan selama penggunaan.

TikTok Shop sebagai *platform social commerce* yang diluncurkan pada tahun 2022 menghadirkan pengalaman belanja terintegrasi dalam aplikasi media sosial [3]. Integrasi antara konten hiburan dan aktivitas transaksi menciptakan pengalaman yang lebih interaktif dibandingkan *e-commerce* konvensional [4]. Pengguna *platform* ini didominasi oleh Generasi Z yang memiliki karakteristik sebagai *digital natives* dengan ekspektasi tinggi terhadap kemudahan penggunaan, kecepatan, dan interaktivitas sistem [5][6]. Namun, tingginya tingkat adopsi tidak selalu diikuti oleh stabilitas retensi pengguna. Kondisi ini menunjukkan adanya kesenjangan antara pengalaman penggunaan yang dirasakan dengan keputusan untuk melanjutkan penggunaan *platform*.

Dalam literatur sistem informasi, *usability* diakui sebagai salah satu determinan utama dalam membentuk pengalaman pengguna [7]. Namun, hubungan antara *usability* dan retensi tidak selalu bersifat langsung. Sistem yang mudah digunakan belum tentu mampu mempertahankan penggunaan apabila pengalaman tersebut tidak memberikan nilai yang cukup untuk penggunaan berkelanjutan [2]. Hal ini menunjukkan bahwa evaluasi *usability* perlu dilakukan secara lebih kontekstual, khususnya pada fitur-fitur yang secara langsung membentuk pengalaman pengguna dalam *platform social commerce*. Dengan demikian, *usability* tidak dapat dipahami semata sebagai atribut teknis sistem, tetapi pengalaman interaksi yang terfragmentasi pada level fitur yang secara kumulatif membentuk evaluasi pengguna terhadap sistem secara keseluruhan.

Pada TikTok Shop, pengalaman pengguna secara dominan dibentuk oleh fitur *live shopping* dan *product reviews*. Fitur *live shopping* memungkinkan interaksi *real-time* antara penjual dan pembeli yang dapat meningkatkan keterlibatan pengguna [8], sementara *product reviews* menyediakan informasi berbasis pengalaman pengguna lain yang memengaruhi persepsi kualitas dan keputusan pembelian [9]. Kedua fitur ini memiliki karakteristik interaksi yang berbeda, sehingga berpotensi menghasilkan pengalaman penggunaan yang berbeda pula. Namun, penelitian sebelumnya cenderung mengevaluasi TikTok Shop secara agregat tanpa membedakan kontribusi masing-masing fitur terhadap pengalaman pengguna [3][10]. Pendekatan agregat tersebut berpotensi mengaburkan mekanisme pengalaman pengguna karena mengasumsikan homogenitas interaksi, padahal fitur dalam *platform social commerce* memiliki karakteristik interaksi yang berbeda secara fundamental. Akibatnya, mekanisme bagaimana pengalaman penggunaan pada level fitur memengaruhi retensi pengguna belum dijelaskan secara memadai.

Selain itu, studi retensi pengguna umumnya menggunakan pendekatan statistik untuk menjelaskan hubungan antar variabel. Pendekatan ini efektif untuk mengidentifikasi hubungan kausal, namun memiliki keterbatasan dalam mengidentifikasi pola perilaku pengguna secara prediktif. Di sisi lain, pendekatan *machine learning* mampu mengidentifikasi pola klasifikasi, tetapi sering kali tidak digunakan secara terpisah dengan pendekatan eksplanatori [11]. Keterpisahan ini menyebabkan analisis retensi pengguna menjadi parsial, di mana hubungan kausal dan pola perilaku dianalisis secara terpisah tanpa keterkaitan kuat dengan kerangka teoretis. Dalam konteks ini, algoritma Naïve Bayes dipilih karena efisien untuk data berukuran kecil, mampu mengolah atribut numerik, serta menghasilkan probabilitas klasifikasi yang sesuai untuk mengidentifikasi pola retensi sebagai pelengkap analisis eksplanatori.

Berdasarkan uraian tersebut, terdapat dua kesenjangan utama dalam penelitian ini. Pertama, terbatasnya kajian yang menganalisis pengaruh *usability* pada level fitur terhadap retensi pengguna dalam konteks *social commerce*. Kedua, belum optimalnya penggabungan pendekatan eksplanatori dan prediktif secara komplementer dalam analisis retensi pengguna. Oleh karena itu, penelitian ini

<http://sistemasi.ftik.unisi.ac.id>

bertujuan untuk menganalisis pengaruh *usability* fitur *live shopping* dan *product reviews* terhadap retensi pengguna TikTok Shop pada Generasi Z, serta memvalidasi pola retensi tersebut menggunakan klasifikasi berbasis Naïve Bayes. Secara teoretis, penelitian ini memperluas pemahaman tentang hubungan antara *usability* dan retensi pengguna dengan menempatkan pengalaman pengguna pada level fitur sebagai unit analisis. Secara metodologis, penelitian ini menggunakan pendekatan eksplanatori dan prediktif secara komplementer dalam satu kerangka analisis, sehingga memungkinkan pemahaman yang lebih komprehensif terhadap perilaku retensi pengguna dalam konteks *social commerce*.

2 Tinjauan Literatur

Penelitian mengenai retensi pengguna dalam sistem informasi secara konsisten menunjukkan bahwa keberlanjutan penggunaan dipengaruhi oleh evaluasi pengalaman pasca penggunaan, yang dimediasi oleh kepuasan dan persepsi nilai. Studi berbasis *Expectation Confirmation Model* menemukan bahwa *confirmation*, *perceived usefulness*, dan kepuasan merupakan determinan utama dalam menjelaskan retensi pengguna pada berbagai konteks digital [2][12][13]. Temuan empiris di Indonesia juga menunjukkan bahwa pengalaman penggunaan yang memenuhi ekspektasi berpengaruh signifikan terhadap kepuasan dan pada akhirnya meningkatkan niat penggunaan berkelanjutan [14]. Namun demikian, hasil penelitian lain menunjukkan bahwa hubungan antara kemudahan penggunaan (*ease of use*) dan retensi tidak selalu signifikan, sehingga mengindikasikan bahwa pengalaman pengguna tidak secara otomatis diterjemahkan menjadi keberlanjutan penggunaan [15]. Dengan demikian meskipun *Expectation Confirmation Model* mampu menjelaskan mekanisme retensi berbasis evaluasi pasca penggunaan, model ini belum secara eksplisit menangkap bagaimana kualitas interaksi pada level fitur membentuk pengalaman yang dievaluasi oleh pengguna.

Di sisi lain, penelitian mengenai *usability* dalam sistem digital umumnya menunjukkan bahwa persepsi kemudahan penggunaan berkontribusi terhadap kepuasan dan niat penggunaan ulang. Studi pada *platform* digital dan layanan berbasis aplikasi menunjukkan bahwa peningkatan kualitas interaksi pengguna dapat meningkatkan kepuasan dan loyalitas pengguna [16]. Namun, sebagian besar penelitian masih mengoperasionalkan *usability* sebagai konstruk agregat pada level sistem [3][7], tanpa mempertimbangkan bahwa pengalaman pengguna dibentuk melalui interaksi terhadap fitur yang berbeda. Penelitian yang mengkaji fitur spesifik, seperti *live streaming* dalam *e-commerce*, menunjukkan bahwa karakteristik interaksi fitur dapat memengaruhi persepsi nilai dan keterlibatan pengguna secara berbeda [17]. Temuan ini mengindikasikan bahwa pendekatan agregat berpotensi menyederhanakan mekanisme pengalaman pengguna, sehingga tidak mampu menjelaskan variasi retensi secara memadai. Oleh karena itu, analisis *usability* yang hanya dilakukan pada level sistem berisiko mengaburkan kontribusi masing-masing fitur dalam membentuk pengalaman pengguna, khususnya pada *platform social commerce* yang mengandalkan interaksi multi-fitur secara simultan.

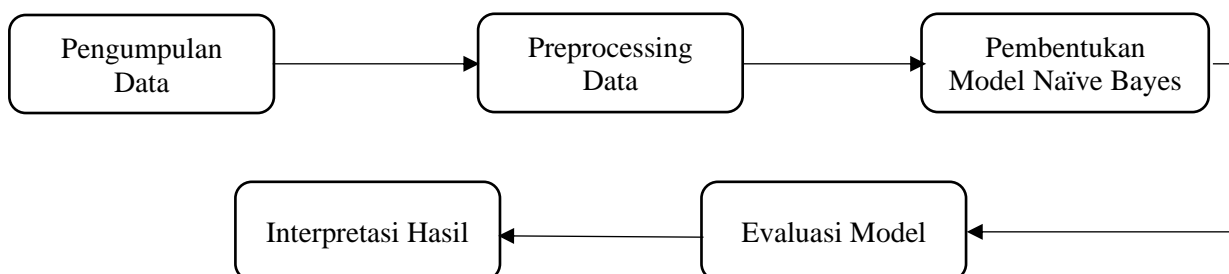
Selain keterbatasan pada level konseptual, penelitian retensi pengguna juga menunjukkan keterbatasan pada pendekatan metodologis. Studi sistem informasi secara dominan menggunakan pendekatan eksplanatori untuk menguji hubungan antar variabel, yang berfokus pada validasi model teoritis [11]. Meskipun pendekatan ini efektif dalam menjelaskan hubungan kausal, penelitian menunjukkan bahwa pendekatan tersebut tidak selalu mampu menangkap pola perilaku pengguna yang kompleks dan dinamis. Di sisi lain, pendekatan berbasis *machine learning* telah digunakan untuk mengidentifikasi pola retensi dan klasifikasi perilaku pengguna dalam berbagai konteks digital [1], namun sering kali digunakan tanpa keterkaitan yang jelas dengan kerangka teoretis.

Dalam konteks klasifikasi perilaku pengguna, salah satu tantangan utama adalah ketidakseimbangan kelas (*class imbalance*), di mana jumlah pengguna yang bertahan dan tidak bertahan tidak proporsional. Kondisi ini menyebabkan model klasifikasi cenderung bias terhadap kelas mayoritas dan menghasilkan metrik evaluasi yang menyesatkan jika hanya mengandalkan akurasi [18][19][20]. Oleh karena itu, evaluasi model prediktif dalam studi retensi perlu mempertimbangkan distribusi data serta keterbatasan metrik klasifikasi dalam merepresentasikan kinerja model secara menyeluruh. Keterbatasan ini memperkuat bahwa pemahaman retensi pengguna menjadi parsial ketika hubungan kausal dan pola perilaku dianalisis secara terpisah. Oleh karena itu, penelitian ini mengombinasikan pendekatan eksplanatori dan prediktif dalam analisis retensi

pengguna serta menempatkan *usability* berbasis fitur sebagai unit analisis untuk memperoleh pemahaman yang lebih komprehensif.

3 Metode Penelitian

Penelitian ini menggunakan pendekatan kuantitatif dengan desain eksplanatori-prediktif untuk menganalisis pengaruh *usability* terhadap retensi pengguna serta mengidentifikasi pola retensi berdasarkan karakteristik pengalaman penggunaan. Pendekatan eksplanatori digunakan untuk menguji hubungan antara *usability* dan retensi pengguna, sedangkan pendekatan prediktif digunakan untuk mengklasifikasikan kecenderungan retensi berdasarkan data empiris [11]. Data yang digunakan merupakan data primer yang dikumpulkan melalui kuesioner digital kepada pengguna TikTok Shop. Populasi penelitian adalah pengguna aktif dari Generasi Z di Indonesia yang telah menggunakan fitur *live shopping* dan *product reviews*. Karena ukuran populasi tidak diketahui secara pasti, jumlah sampel ditentukan menggunakan pendekatan Hair et al. [21], yaitu 5-10 kali jumlah indikator. Dengan total 23 indikator, diperoleh jumlah minimum 115 responden. Teknik pengambilan sampel menggunakan purposive sampling dengan kriteria: (1) berusia 13-28 tahun, (2) aktif menggunakan TikTok Shop, dan (3) pernah menggunakan kedua fitur yang diteliti. Selain jumlah responden, distribusi kelas target dalam penelitian ini juga dianalisis untuk memahami karakteristik data yang digunakan dalam proses klasifikasi. Berdasarkan hasil pengelompokan data retensi pengguna, jumlah responden yang termasuk dalam kategori “Ya” (retensi) sebanyak 101 responden, sedangkan kategori “Tidak” (tidak retensi) sebanyak 14 responden. Distribusi ini menimbulkan kondisi ketidakseimbangan kelas, di mana proporsi kelas “Ya” (87,8 persen) jauh lebih dominan dibandingkan kelas “Tidak” (12,2 persen), sehingga berpotensi memengaruhi kinerja model klasifikasi, khususnya dalam mendeteksi kelas minoritas.



Gambar 1 Alur penelitian

Gambar 1 menunjukkan alur penelitian ini. Tahapan dimulai dari pengumpulan data menggunakan kuesioner *online*, dilanjutkan dengan *preprocessing* data yang meliputi pembersihan data, pengkodean variabel, dan pembentukan dataset klasifikasi sebelum dilakukan pemodelan menggunakan algoritma Naïve Bayes. Selanjutnya, model yang dihasilkan dievaluasi menggunakan metrik *accuracy*, *precision*, *recall*, dan *F1-score*, kemudian dilakukan interpretasi hasil untuk memperoleh kesimpulan penelitian.

Instrumen Penelitian

Instrumen penelitian menggunakan *System Usability Scale* (SUS) untuk mengukur persepsi *usability* pada fitur *live shopping* dan *product reviews*. Setiap fitur diukur menggunakan 10 item pernyataan dengan skala *Likert* 5 poin, mulai dari sangat tidak setuju (1) hingga sangat setuju (5). Pengukuran dilakukan secara terpisah untuk masing-masing fitur sehingga setiap responden menghasilkan dua skor *usability*. Daftar item pernyataan yang digunakan dalam pengukuran *usability* berdasarkan SUS yang dikembangkan Brooke [22], seperti yang disajikan pada Tabel 1.

Tabel 1 Instrumen System Usability Scale (SUS) [22]

Item	Pertanyaan
Q1	Saya berpikir akan menggunakan fitur ini lagi.
Q2	Saya merasa fitur ini rumit untuk digunakan.
Q3	Saya merasa fitur ini mudah untuk digunakan.
Q4	Saya membutuhkan bantuan dari orang lain atau teknisi dalam menggunakan fitur ini.
Q5	Saya merasa fitur dalam aplikasi ini berjalan dengan semestinya.

-
- Q6 Saya merasa ada banyak hal yang tidak konsisten (tidak serasi) pada aplikasi ini.
Q7 Saya merasa orang lain akan memahami cara menggunakan fitur ini dengan cepat.
Q8 Saya merasa fitur ini membingungkan.
Q9 Saya merasa tidak ada hambatan dalam menggunakan fitur ini.
Q10 Saya perlu membiasakan diri terlebih dahulu sebelum menggunakan fitur ini.
-

Pengolahan skor *usability* dilakukan menggunakan metode SUS yang menghasilkan skor dalam rentang 0-100. Perhitungan dilakukan dengan mengurangi skor pada item ganjil dengan 1, serta mengurangi nilai 5 dengan skor pada item genap. Seluruh skor kemudian dijumlahkan dan dikalikan dengan faktor 2,5 untuk memperoleh skor akhir *usability*. Interpretasi skor SUS mengacu pada Bangor et al. [23], di mana skor ≥ 70 dikategorikan sebagai dapat diterima (acceptable), skor 70-80 sebagai baik, dan skor > 80 sebagai sangat baik. Skor ini digunakan sebagai representasi kuantitatif *usability* dalam analisis eksplanatori maupun prediktif.

Selain itu, variabel retensi pengguna diukur menggunakan indikator perilaku yang mencerminkan keberlanjutan penggunaan, yaitu frekuensi penggunaan ulang, intensi penggunaan, dan durasi penggunaan. Ketiga indikator ini digunakan untuk merepresentasikan dimensi perilaku dan niat dalam *continuance usage*, sebagaimana umum digunakan dalam studi sistem informasi [2][24][25]. Dalam penelitian ini, retensi pengguna diklasifikasikan sebagai “Ya” apabila responden memenuhi dua dari tiga kriteria berikut: (1) frekuensi penggunaan ≥ 3 kali per bulan, (2) memiliki intensi untuk menggunakan kembali, dan (3) durasi penggunaan ≥ 15 menit per hari. Pendekatan ini digunakan untuk menghindari bias pengukuran berbasis satu indikator tunggal, serta memastikan bahwa retensi mencerminkan kombinasi antara perilaku aktual dan niat penggunaan. Responden yang tidak memenuhi kriteria tersebut diklasifikasikan sebagai “Tidak”.

Teknik Analisis Data

Sebelum analisis utama dilakukan, data diuji menggunakan uji validitas *Pearson Product Moment* dengan kriteria r -hitung $>$ r -tabel pada tingkat signifikansi 0,05, serta uji reliabilitas menggunakan *Cronbach's Alpha* dengan batas minimum 0,70 untuk memastikan konsistensi internal instrumen [21]. Analisis data dilakukan dalam dua tahap yang saling melengkapi. Tahap pertama, analisis eksplanatori untuk menguji pengaruh *usability* terhadap retensi pengguna menggunakan metode analisis regresi logistik biner, karena variabel dependen berbentuk kategorik (Ya/Tidak). Skor SUS digunakan sebagai variabel independen dalam model ini.

Tahap kedua, analisis prediktif menggunakan algoritma Naïve Bayes yang mengklasifikasikan retensi pengguna berdasarkan skor *usability* dari masing-masing fitur yang digunakan sebagai atribut numerik dalam model klasifikasi probabilistik, dengan variabel retensi sebagai kelas target. Algoritma ini dipilih karena mampu menangani data numerik dan kategorikal secara efisien serta menghasilkan probabilitas klasifikasi berbasis teori Bayes [26][27]. Secara matematis, probabilitas posterior suatu kelas dihitung menggunakan Teorema Bayes, sebagai berikut

$$P(C|X) = \frac{P(X|C) \times P(C)}{P(X)}$$

dengan:

$P(C|X)$ = probabilitas posterior kelas (retensi: Ya/Tidak)

$P(X|C)$ = *likelihood* data terhadap kelas

$P(C)$ = probabilitas *prior*

$P(X)$ = probabilitas *evidence*

Dalam asumsi independensi antar atribut, *likelihood* dihitung sebagai:

$$P(x|C) = \prod_{i=1}^n P(x_i | C)$$

di mana x_i merupakan atribut input skor *usability*.

Pendekatan ini memungkinkan identifikasi pola retensi pengguna yang tidak sepenuhnya terjelaskan melalui model eksplanatori berbasis regresi, sehingga memberikan perspektif tambahan dalam

memahami perilaku keberlanjutan penggunaan. Evaluasi model klasifikasi dilakukan menggunakan metode *10-Fold Cross Validation* untuk memastikan stabilitas hasil. Kinerja model diukur menggunakan metrik *accuracy*, *precision*, *recall*, dan *F1-score*. Seluruh proses analisis dilakukan menggunakan perangkat lunak SPSS untuk analisis dan RapidMiner untuk pemodelan klasifikasi.

4 Hasil dan Pembahasan

Hasil Penelitian

Data penelitian diperoleh dari 115 responden Generasi Z yang merupakan pengguna aktif TikTok Shop dan telah menggunakan fitur *live shopping* dan *product reviews*. Mayoritas responden berada pada rentang usia 19-21 tahun dan didominasi oleh perempuan (69,4 persen), dengan sebagian besar berstatus mahasiswa (84,3 persen). Komposisi ini menunjukkan bahwa data merepresentasikan kelompok pengguna *digital native* yang memiliki intensitas interaksi tinggi dengan *platform social commerce*. Hasil uji validitas menunjukkan bahwa seluruh item pada variabel *live shopping* dan *product reviews* memiliki nilai r-hitung yang lebih besar dibandingkan r-tabel (0,183), sehingga seluruh item dinyatakan valid dan layak digunakan dalam analisis lanjutan. Uji reliabilitas menunjukkan nilai *Cronbach's Alpha* sebesar 0,772 untuk *live shopping* dan 0,752 untuk *product reviews*, yang mengindikasikan konsistensi internal yang baik dan layak digunakan dalam analisis lanjutan.

Analisis *usability* menggunakan *System Usability Scale* (SUS) menunjukkan bahwa fitur *live shopping* memperoleh skor rata-rata sebesar 52, sedangkan fitur *product reviews* memperoleh skor rata-rata sebesar 53. Ringkasan hasil perhitungan SUS disajikan pada Tabel 2, yang menunjukkan bahwa kedua fitur berada pada kategori *Marginal-Low*, yaitu di bawah ambang batas kelayakan sistem (≥ 70). Temuan ini mengindikasikan bahwa meskipun kedua fitur dapat digunakan, pengalaman penggunaan yang dihasilkan belum optimal dalam hal kemudahan, efisiensi, dan kenyamanan.

Tabel 2 Ringkasan skor SUS

Fitur	Skor Rata-rata	Kategori
<i>Live shopping</i>	52	<i>Marginal-Low</i>
<i>Product reviews</i>	53	<i>Marginal-Low</i>

Sumber: Data primer diolah (2026)

Hasil regresi logistik biner disajikan pada Tabel 3, menunjukkan bahwa variabel *usability* fitur *live shopping* tidak memiliki pengaruh signifikan terhadap retensi pengguna ($B = -0,028$; $p = 0,481$). Demikian pula, *usability* fitur *product reviews* juga tidak menunjukkan pengaruh signifikan terhadap retensi pengguna ($B = 0,061$; $p = 0,141$). Secara arah, koefisien *live shopping* bernilai negatif dengan *Odds Ratio* sebesar 0,972, yang menunjukkan bahwa peningkatan *usability* pada fitur tersebut cenderung diikuti oleh penurunan peluang retensi, meskipun pengaruh ini sangat lemah dan tidak signifikan. Sebaliknya, fitur *product reviews* memiliki koefisien positif dengan *Odds Ratio* sebesar 1,063, yang menunjukkan kecenderungan peningkatan peluang retensi, namun dengan besaran pengaruh yang kecil dan tidak signifikan. Temuan ini menunjukkan bahwa tidak terdapat bukti empiris yang cukup untuk menyatakan bahwa *usability* pada level fitur berpengaruh terhadap retensi pengguna dalam model yang diuji.

Tabel 3 Hasil regresi logistik biner

Variabel	Koefisien (B)	Std. Error	Wald	Sig. (p-value)	Odds Ratio Exp(B)
<i>Live shopping</i>	-0,028	0,040	0,496	0,481	0,972
<i>Product reviews</i>	0,061	0,041	2,164	0,141	1,063
Konstanta	0,507	2,806	0,033	0,857	1,659

Sumber: Data primer diolah (2026)

Nilai Nagelkerke R^2 sebesar 0,046 di Tabel 4 menunjukkan bahwa model hanya mampu menjelaskan 4,6 persen variasi retensi pengguna, yang menindikasikan bahwa faktor lain di luar *usability* lebih dominan dalam mempengaruhi retensi. Sementara itu, hasil uji Hosmer-Lemeshow dengan nilai signifikansi sebesar 0,507 ($p > 0,05$) menunjukkan bahwa model memiliki kecocokan yang baik dengan data. Secara keseluruhan, hasil ini menunjukkan bahwa model regresi logistik yang dibangun belum mampu menjelaskan retensi pengguna secara kuat berdasarkan variabel *usability* pada level fitur.

Tabel 4 Goodness of fit model

Indikator	Nilai
-----------	-------

Nagelkerke R ²	0,046
Hosmer-Lemeshow Test (Sig.)	0,507

Sumber: Data primer diolah (2026)

Hasil analisis prediktif menggunakan algoritma Naïve Bayes menunjukkan bahwa model klasifikasi menghasilkan nilai *accuracy* sebesar 88,64 persen, namun performa ini tidak merata antar kelas. Distribusi hasil klasifikasi ditunjukkan pada Tabel 5 yang menggambarkan perbandingan antara hasil prediksi model dan kondisi aktual. Model menunjukkan kemampuan yang relatif baik dalam mengidentifikasi pengguna yang bertahan (*True Positive* = 100), namun memiliki keterbatasan signifikan dalam mendeteksi pengguna yang tidak bertahan, di mana sebagian besar salah diklasifikasikan sebagai bertahan. Hal ini mengindikasikan adanya bias model terhadap kelas mayoritas akibat ketidakseimbangan distribusi kelas.

Tabel 5 Confusion matrix

	True Ya	True Tidak
Pred. Ya	100	12
Pred. Tidak	1	2

Sumber: Data primer diolah (2026)

Evaluasi lebih lanjut terhadap kinerja model disajikan pada Tabel 6, yang menunjukkan nilai *precision*, *recall*, dan *F1-score* untuk masing-masing kelas. Model memiliki performa yang sangat tinggi pada kelas “Ya”, namun performa pada kelas “Tidak” relatif rendah, terutama pada nilai *recall* sebesar 14,29 persen. Kondisi ini menunjukkan adanya ketidakseimbangan distribusi kelas (*class imbalance*), yang menyebabkan model lebih optimal dalam mengidentifikasi kelas mayoritas dibandingkan kelas minoritas. Model kurang mampu mendeteksi pengguna yang tidak bertahan (*churn*). Nilai *macro-average F1-score* sebesar 58,71 persen mengindikasikan bahwa kinerja model secara keseluruhan masih dipengaruhi oleh ketidakseimbangan distribusi kelas. Hal ini menunjukkan bahwa model memiliki keterbatasan dalam mengidentifikasi pengguna yang berpotensi *churn*.

Tabel 6 Evaluasi kinerja model

Kelas	<i>Precision</i>	<i>Recall</i>	<i>F1-Score</i>
Ya	89,29%	99,01%	93,89%
Tidak	66,67%	14,29%	23,53%
<i>Macro-average</i>	77,98%	56,65%	58,71%

Sumber: Data primer diolah (2026)

Visualisasi kinerja model pada Gambar 2 hingga Gambar 5 menunjukkan bahwa performa model sangat dominan pada kelas “Ya”, terutama pada nilai *recall* yang mendekati 100 persen, sementara performa pada kelas “Tidak” relatif rendah. Pola ini memperkuat indikasi adanya bias model terhadap kelas mayoritas serta keterbatasan model dalam mendeteksi pengguna yang tidak bertahan.

accuracy: 88.64% +/- 5.89% (micro average: 88.70%)			
	true Ya	true Tidak	class precision
pred Ya	100	12	89.29%
pred Tidak	1	2	66.67%
class recall	99.01%	14.29%	

Gambar 2 Visualisasi kinerja model (*accuracy*)

precision: 66.67% (positive class: Tidak)			
	true Ya	true Tidak	class precision
pred Ya	100	12	89.29%
pred Tidak	1	2	66.67%
class recall	99.01%	14.29%	

Gambar 3 Visualisasi kinerja model (*precision*)

Gambar 4 Visualisasi kinerja model (*recall*)

recall: 15.00% +/- 33.75% (micro average: 14.29%) (positive class: Tidak)			
	true Ya	true Tidak	class precision
pred Ya	100	12	89.29%
pred Tidak	1	2	66.67%
class recall	99.01%	14.29%	

Gambar 5 Visualisasi kinerja model (*F1-Score*)

Pembahasan

Temuan penelitian menunjukkan bahwa *usability* pada fitur *live shopping* dan *product reviews* berada pada kategori *marginal-low*, dengan skor masing-masing 52 dan 53. Berdasarkan interpretasi *System Usability Scale* (SUS), skor di bawah ambang batas 68-70 menunjukkan bahwa sistem masih berada pada tingkat penerimaan yang rendah meskipun tetap dapat digunakan secara fungsional [22][23]. Hal ini menunjukkan bahwa keberfungsian teknis suatu fitur tidak selalu diikuti oleh kualitas pengalaman yang memadai untuk memperkuat evaluasi positif pengguna terhadap sistem.

Namun demikian, hasil regresi logistik menunjukkan bahwa *usability* pada fitur *live shopping* dan *product reviews* tidak memiliki pengaruh signifikan terhadap retensi pengguna. Temuan ini menunjukkan bahwa dalam konteks *social commerce*, *usability* pada level fitur memiliki keterbatasan dalam menjelaskan retensi pengguna. Meskipun secara konseptual *usability* sering diposisikan sebagai determinan penting dalam adopsi dan keberlanjutan penggunaan sistem, hasil penelitian ini tidak memberikan bukti empiris yang cukup untuk mendukung keberadaan hubungan tersebut pada level analisis fitur. Hal ini mengindikasikan bahwa hubungan antara *usability* dan retensi bersifat kontekstual dan bergantung pada karakteristik sistem yang digunakan, di mana pada *platform* yang bersifat interaktif dan berbasis konten, keputusan pengguna untuk bertahan tidak semata ditentukan oleh kemudahan penggunaan, tetapi oleh pengalaman yang lebih luas pada level *platform*. Dengan demikian, temuan ini tidak menolak peran *usability* secara keseluruhan, tetapi menunjukkan bahwa pengaruhnya tidak selalu bersifat langsung dan dominan dalam konteks *social commerce*. Hasil ini konsisten dengan *Expectation Confirmation Model* yang menyatakan bahwa retensi tidak hanya ditentukan oleh pengalaman awal penggunaan, tetapi oleh evaluasi pasca penggunaan yang mencakup kepuasan dan persepsi nilai [2]. Selain itu, penelitian empiris juga menunjukkan bahwa kemudahan penggunaan tidak selalu berpengaruh signifikan terhadap retensi pengguna, terutama pada sistem yang melibatkan interaksi kompleks [15]. Ketiadaan pengaruh *usability* dalam penelitian ini mengindikasikan bahwa retensi pengguna kemungkinan dipengaruhi oleh faktor lain di luar *usability*, seperti keterlibatan berbasis konten dan interaksi sosial yang lebih dominan dalam ekosistem *social commerce* [28][29].

Di sisi lain, hasil analisis prediktif menggunakan Naïve Bayes menunjukkan bahwa model mampu mencapai *accuracy* sebesar 88,64 persen, namun kinerja ini tidak merata antar kelas. Model menunjukkan kemampuan yang baik dalam mengidentifikasi pengguna yang bertahan, namun memiliki keterbatasan signifikan dalam mendeteksi pengguna yang tidak bertahan. Kondisi ini menunjukkan bahwa model lebih dapat diandalkan untuk mengenali pola pengguna yang bertahan, namun kurang memadai untuk mengidentifikasi risiko *churn*. Hal ini tercermin dari rendahnya nilai *recall* kelas “Tidak” dan menunjukkan adanya ketidakseimbangan kelas, di mana model cenderung bias terhadap kelas mayoritas [19][20]. Dalam kondisi ini, akurasi tidak selalu menjadi indikator yang representatif terhadap kinerja model secara keseluruhan [30]. Hal ini menunjukkan bahwa *accuracy* sebagai metrik tunggal tidak cukup untuk mengevaluasi kinerja model klasifikasi pada data yang tidak seimbang. Selanjutnya, perilaku pengguna yang bertahan relatif lebih konsisten dan dapat dikenali, sedangkan perilaku pengguna yang tidak bertahan lebih heterogen dan tidak cukup dijelaskan oleh atribut *usability* yang digunakan dalam model, yang tercermin dari rendahnya nilai *recall* pada kelas “Tidak”. Hal ini menegaskan pentingnya penggunaan metrik evaluasi lain seperti *precision*, *recall*, dan *F1-score* untuk memperoleh gambaran performa model yang lebih komprehensif [18].

Perbedaan hasil antara regresi logistik dan Naïve Bayes menunjukkan bahwa retensi pengguna tidak sepenuhnya dapat dijelaskan melalui hubungan kausal linier, namun menunjukkan kecenderungan pola probabilistik yang dapat dikenali secara terbatas. Regresi logistik berfungsi untuk menguji signifikansi hubungan antar variabel, sedangkan Naïve Bayes mengidentifikasi pola distribusi data untuk keperluan klasifikasi. Perbedaan ini sejalan dengan literatur yang membedakan pendekatan eksplanatori dan prediktif dalam penelitian sistem informasi [11]. Secara keseluruhan, temuan penelitian ini menunjukkan bahwa *usability* berbasis fitur tidak terbukti secara signifikan sebagai determinan retensi pengguna. Selain itu, kombinasi pendekatan eksplanatori dan prediktif memberikan pemahaman yang lebih komprehensif terhadap perilaku retensi, khususnya dalam membedakan antara hubungan kausal dan pola perilaku pengguna. Temuan ini mengindikasikan kemungkinan bahwa retensi pengguna dipengaruhi oleh faktor lain di luar *usability*, termasuk keterlibatan pengguna dan dinamika interaksi sosial dalam ekosistem *platform*.

Temuan tersebut sejalan dengan penelitian sebelumnya yang menunjukkan bahwa performa terbaik pada algoritma Naïve Bayes justru diperoleh dari model *default* tanpa penerapan teknik penanganan ketidakseimbangan data [20][30]. Pada algoritma probabilistik seperti Naïve Bayes, penggunaan data sintetis melalui SMOTE berpotensi mengubah distribusi probabilitas fitur dan mengganggu asumsi independensi antar fitur, sehingga dapat menurunkan performa model [20]. Selain itu, metode *Random Undersampling* juga berisiko menghilangkan sebagian informasi penting dari kelas mayoritas [20]. Oleh karena itu, dalam penelitian ini model Naïve Bayes digunakan pada distribusi data asli tanpa penerapan teknik *balancing*, karena pendekatan tersebut masih relevan dan didukung oleh penelitian sebelumnya yang menunjukkan bahwa Naïve Bayes dapat bekerja optimal tanpa modifikasi distribusi kelas [30].

5 Kesimpulan

Penelitian ini bertujuan untuk menganalisis pengaruh *usability* fitur *live shopping* dan *product reviews* terhadap retensi pengguna TikTok Shop pada Generasi Z, serta mengidentifikasi pola retensi menggunakan pendekatan prediktif. Hasil penelitian menunjukkan bahwa tingkat *usability* berada pada kategori *marginal-low*. Selain itu, hasil analisis regresi logistik menunjukkan bahwa *usability* tidak berpengaruh signifikan terhadap retensi pengguna. Di sisi lain, hasil klasifikasi menggunakan Naïve Bayes menunjukkan bahwa model memiliki tingkat akurasi yang tinggi sebesar 88,64 persen dalam klasifikasi retensi, namun memiliki keterbatasan dalam mendeteksi pengguna yang tidak bertahan akibat ketidakseimbangan distribusi kelas. Temuan ini menunjukkan bahwa pendekatan prediktif menggunakan algoritma Naïve Bayes mampu memberikan kontribusi dalam mengidentifikasi pola retensi pengguna secara sistematis. Dengan demikian, retensi pengguna dalam konteks *social commerce* tidak sepenuhnya dijelaskan oleh faktor *usability*, serta menunjukkan bahwa pendekatan eksplanatori dan prediktif memberikan hasil yang berbeda namun, saling melengkapi dalam memahami perilaku pengguna. Penelitian ini memiliki beberapa keterbatasan. Pertama, variabel yang digunakan dalam penelitian ini terbatas pada aspek *usability* berbasis fitur, sehingga belum mampu menangkap faktor lain yang berpotensi memengaruhi retensi pengguna, seperti kepercayaan, kepuasan, maupun keterlibatan pengguna dalam ekosistem *platform*. Kedua, distribusi data yang tidak seimbang antara pengguna yang bertahan dan tidak bertahan memengaruhi kinerja model klasifikasi, khususnya dalam mendeteksi kelas minoritas. Ketiga, penelitian ini menggunakan pendekatan *cross-sectional* sehingga belum mampu menggambarkan dinamika perilaku retensi pengguna secara longitudinal. Oleh karena itu, penelitian selanjutnya disarankan untuk memperluas model dengan memasukkan variabel lain yang lebih komprehensif, mengembangkan desain longitudinal, serta menerapkan teknik penanganan ketidakseimbangan kelas. Secara lebih spesifik, penelitian mendatang dapat mengeksplorasi penggunaan algoritma berbasis *ensemble* seperti Random Forest atau XGBoost, yang lebih mampu menangkap pola non-linear serta mengurangi bias terhadap kelas mayoritas pada data yang tidak seimbang. Pendekatan ini relevan mengingat temuan penelitian menunjukkan keterbatasan Naïve Bayes dalam mendeteksi pengguna yang tidak bertahan, sehingga penggunaan algoritma *ensemble* berpotensi meningkatkan sensitivitas terhadap kelas minoritas serta menghasilkan klasifikasi yang lebih seimbang dalam konteks retensi pengguna.

Referensi

- [1] W. F. Mustafa, S. Hidayat, and D. H. Fudholi, "Prediksi Retensi Pengguna Baru Shopee menggunakan Machine Learning," *J. Media Inform. Budidarma*, Vol. 8, No. 1, pp. 612–623, 2024, DOI: 10.30865/mib.v8i1.7074.
- [2] A. Bhattacharjee, "Understanding Information Systems Continuance: An Expectation-Confirmation Model," *MIS Q.*, Vol. 25, No. 3, pp. 351–370, Sep. 2001, DOI: 10.2307/3250921.
- [3] F. M. Tambajong, R. Sengkey, and S. D. E. Paturusi, "Analyzing TikTok Shop User Experience with System Usability Scale Method," *J. Tek. Inform.*, Vol. 19, No. 3, pp. 251–258, 2024, DOI: 10.35793/jti.v19i3.53811.
- [4] F. Fieryando and B. Kristianto, "Analisis Sentimen terhadap TikTok Shop dengan K-Nearest Neighbor, Decision Tree, dan Naive Bayes," *J. Buana Inform.*, Vol. 15, No. 1, pp. 21–29, 2024, DOI: 10.24002/jbi.v15i1.8205.
- [5] DataReportal, "Essential TikTok Stats," DataReportal.
- [6] F. Fathiyah, "TikTok dan kebebasan berekspresi di Ruang Digital bagi Generasi Z," *MEDIALOG J. Ilmu Komun.*, Vol. 6, No. 1, pp. 166–177, 2023, DOI: 10.35326/medialog.v6i1.4139.
- [7] I. Purwandani, N. O. Syamsiah, and S. Nurwahyuni, "Perceived Usability Evaluation of TikTok Shop Platform using the System Usability Scale," *Sink. J. dan Penelit. Tek. Inform.*, Vol. 7, No. 3, pp. 1389–1399, 2023, DOI: 10.33395/sinkron.v8i3.12473.
- [8] I. Zuhdi, L. M. Furkan, and H. Hilmiahi, "Pengaruh Live Streaming Shopping dan Motivasi Belanja Hedonis terhadap Impulsif Buying pada Pengguna Aplikasi TikTok Shop di Kota Mataram," *J. Ris. Pemasar.*, Vol. 2, No. 3, pp. 66–72, 2023, DOI: 10.29303/jrp.v2i3.3286.
- [9] R. C. Kawet, I. D. Palandeng, and R. V. Pitta, "Pengaruh Online Customer Review, Online Customer Rating, dan Cash on Delivery terhadap Keputusan Pembelian Produk pada E-Commerce Tokopedia," *JMBI UNSRAT (Jurnal Ilm. Manaj. Bisnis dan Inov. Univ. Sam Ratulangi)*, Vol. 11, No. 1, pp. 1170–1183, 2024, DOI: 10.35794/jmbi.v11i1.56516.
- [10] I. W. Sanjaya, M. R. Indah, D. R. Kurniawan, D. Saputra, and Ariansyah, "Analisis Tingkat Kepuasan Pengguna Layanan Tiktok Shop menggunakan System Usability Scale (SUS)," *J. JUPITER*, Vol. 15, No. Oktober 2023, pp. 951–959, 2021, DOI: 10.5281/zenodo.10068748.
- [11] G. Shmueli and O. R. Koppius, "Predictive Analytics in Information Systems Research," *MIS Q.*, Vol. 35, No. 3, pp. 553–572, Sep. 2011, DOI: 10.2307/23042796.
- [12] G.-D. Nguyen and T.-H. T. Dao, "Factors Influencing Continuance Intention to use Mobile Banking: An Extended Expectation-Confirmation Model with Moderating Role of Trust," *Humanit. Soc. SCI. Commun.*, Vol. 11, No. 1, p. 276, Feb. 2024, DOI: 10.1057/s41599-024-02778-z.
- [13] A. A. Daneji, A. F. M. Ayub, and M. N. M. Khambari, "The Effects of Perceived Usefulness, Confirmation and Satisfaction on Continuance Intention in using Massive Open Online Course (MOOC)," *Knowl. Manag. E-Learning An Int. J.*, Vol. 11, No. 2, pp. 201–214, Jun. 2019, DOI: 10.34105/j.kmel.2019.11.010.
- [14] R. Alfiyaturohmah, M. Prabantarikso, and B. M. Simatupang, "The Effect of Expectation Confirmation and Perceived Usefulness on Continuance Intention of QRIS Mobile Banking Users: The Mediating Role of Customer Satisfaction," *KALBISOCIO J. Bisnis dan Komun.*, Vol. 13, No. 1, pp. 7–19, 2026.
- [15] A. A. M Ayyoub, B. A. Abu Eidah, Z. N. Khlaif, M. Ahmad EL-Shamali, and M. R. Sulaiman, "Understanding Online Assessment Continuance Intention and Individual Performance by Integrating Task Technology Fit and Expectancy Confirmation Theory," *Heliyon*, Vol. 9, No. 11, p. e22068, Nov. 2023, DOI: 10.1016/j.heliyon.2023.e22068.
- [16] F. M. Asyraf, A. Anofrizen, S. Syaifullah, and E. Saputra, "The Impact of User Satisfaction on the Intention to Continuously use the Threads Application: An ECM approach," *SISTEMASI*, Vol. 14, No. 2, pp. 674–686, Mar. 2025, DOI: 10.32520/stmsi.v14i2.5027.
- [17] I. Z. Abdilllah and H. N. Fadhillah, "Pengaruh Perceived Usefulness, Confirmation, dan Satisfaction terhadap Continuance Intention Penggunaan Fitur Live Streaming Shopee Eiger," *J. Manaj. Terap. dan Keuang.*, Vol. 15, No. 1, 2026, DOI:

- <https://doi.org/10.22437/jmk.v15i01.52916>.
- [18] M. Sokolova and G. Lapalme, "A Systematic Analysis of Performance Measures for Classification Tasks," *Inf. Process. Manag.*, Vol. 45, No. 4, pp. 427–437, Jul. 2009, DOI: 10.1016/j.ipm.2009.03.002.
- [19] Haibo He and E. A. Garcia, "Learning from Imbalanced Data," *IEEE Trans. Knowl. Data Eng.*, Vol. 21, No. 9, pp. 1263–1284, Sep. 2009, DOI: 10.1109/TKDE.2008.239.
- [20] F. M. Megahed, Y.-J. Chen, A. Megahed, Y. Ong, N. Altman, and M. Krzywinski, "The Class Imbalance Problem," *Nat. Methods*, Vol. 18, No. 11, pp. 1270–1272, Nov. 2021, DOI: 10.1038/s41592-021-01302-4.
- [21] J. F. Hair, B. J. Babin, R. E. Anderson, and W. C. Black, *Multivariate Data Analysis*, 8th Editio. United Kingdom: Annabel Ainscow, 2018.
- [22] J. Brooke, "SUS: A Quick and Dirty Usability Scale," in *Usability Evaluation in Industry*, P. W. Jordan, B. Thomas, B. A. Weerdmeester, and I. L. McClelland, Eds., Taylor & Francis, 1996, pp. 189–194. DOI: 10.1201/9781498710411-35.
- [23] A. Bangor, P. Kortum, and J. Miller, "Determining What Individual SUS Scores Mean: Adding an Adjective Rating Scale," *JUS J. Usability Stud.*, Vol. 4, No. 3, pp. 114–123, 2009.
- [24] W. H. DeLone and E. R. McLean, "The DeLone and McLean Model of Information Systems Success: A Ten-Year Update," *J. Manag. Inf. Syst.*, Vol. 19, No. 4, pp. 9–30, 2003, DOI: 10.1080/07421222.2003.11045748.
- [25] V. Venkatesh, M. G. Morris, G. B. Davis, and F. D. Davis, "User Acceptance of Information Technology: Toward a Unified View," *MIS Q. Manag. Inf. Syst.*, Vol. 27, No. 3, pp. 425–478, 2003, DOI: 10.2307/30036540.
- [26] M. D. Kurnia, T. Tukiyat, and M. Makhsun, "Penerapan Metode Naïve Bayes dan Weighted Product untuk Prediksi Lanjut Studi Peserta Didik," *J. Teknol. Sist. Inf. dan Apl.*, Vol. 6, No. 4, pp. 730–739, 2023, DOI: 10.32493/jtsi.v6i4.33538.
- [27] S. S. Hasibuan, A. Angraini, E. Saputra, and M. Megawati, "Sentimen Analisis terhadap Fitur TikTok Shop menggunakan Naïve Bayes dan K-Nearest Neighbor," *J. Media Inform. Budidarma*, Vol. 8, No. 1, pp. 303–311, 2024, DOI: 10.30865/mib.v8i1.7238.
- [28] N. Hajli, "Social Commerce Constructs and Consumer's Intention to Buy," *Int. J. Inf. Manage.*, Vol. 35, No. 2, pp. 183–191, Apr. 2015, DOI: 10.1016/j.ijinfomgt.2014.12.005.
- [29] K. Z. K. Zhang and M. Benyoucef, "Consumer Behavior in Social Commerce: A Literature Review," *Decis. Support Syst.*, Vol. 86, pp. 95–108, Jun. 2016, DOI: 10.1016/j.dss.2016.04.001.
- [30] N. Japkowicz and S. Stephen, "The Class Imbalance Problem: A Systematic Study," *Intell. Data Anal.*, Vol. 6, No. 5, pp. 429–449, 2002.