PENGEMBANGAN MODEL XAI BERDASARKAN PERSEPSI PENGGUNA UNTUK PENILAIAN RISIKO KREDIT PADA P2P LENDING DI INDONESIA

Yefta Christian, Jesen Jeverlino, Tony Wibowo
Sistem Informasi, Universitas Internasional Batam
Jl. Gajah Mada, Tiban Indah, Kec. Sekupang, Kota Batam, Indonesia
yefta@uib.ac.id

ABSTRAK

Pertumbuhan layanan *peer-to-peer* (P2P) *lending* di Indonesia meningkatkan kebutuhan akan penilaian risiko kredit yang akurat dan transparan. Penggunaan *explainable artificial intelligence* (XAI) menjadi salah satu solusi atas keterbatasan model kecerdasan buatan konvensional yang bersifat *black-box*. Meskipun XAI dirancang untuk meningkatkan transparansi dan pemahaman model, kajian mengenai hubungan antara XAI dan persepsi pengguna, khususnya faktor psikologis yang memengaruhi kepercayaan dan keputusan investasi, masih terbatas. Oleh karena itu, penelitian ini bertujuan untuk mengembangkan model XAI yang mempertimbangkan persepsi pengguna dalam penilaian risiko kredit pada P2P *lending* di Indonesia. Pendekatan kuantitatif digunakan melalui survei terhadap 377 responden yang dianalisis menggunakan PLS-SEM untuk mengidentifikasi faktor-faktor yang memengaruhi niat adopsi XAI. Hasil analisis tersebut menjadi dasar pengembangan model XAI menggunakan kerangka CRISP-DM. Hasil penelitian menunjukkan bahwa *perceived value* merupakan faktor terkuat dalam membentuk niat adopsi, dengan *perceived need* dan AI *self-efficacy* berpengaruh signifikan terhadap *perceived value*, sementara AI *anxiety* hanya berpengaruh terhadap *perceived need*. Temuan ini kemudian digunakan sebagai dasar perancangan model XAI yang mampu memberikan penilaian risiko kredit secara akurat dan dapat dijelaskan sesuai persepsi pengguna melalui interpretasi SHAP, dengan kinerja tinggi (AUC 0,935) setelah penanganan ketidakseimbangan data dan optimasi *hyperparameter* menggunakan Optuna.

Kata kunci : *P2P Lending, Penilaian Risiko Kredit, Explainable Artificial Intelligence, Persepsi Pengguna*

1. PENDAHULUAN

Perkembangan teknologi yang pesat telah melahirkan berbagai inovasi di sektor keuangan, salah satunya adalah *peer-to-peer* (P2P) *lending* yang kini menjadi alternatif layanan pinjaman populer di Indonesia [1]. Di balik pertumbuhannya, P2P *lending* sering menghadapi risiko kredit karena sejumlah debitur tidak mampu memenuhi kewajibannya dalam membayar pinjaman. Kredit yang diberikan tanpa mempertimbangkan kualitas debitur meningkatkan risiko gagal bayar dan berpotensi menjadi bermasalah, karena dapat menurunkan likuiditas dan menambah kebutuhan cadangan dana untuk menutup kerugian. Oleh sebab itu, menjaga kualitas kredit menjadi hal yang penting untuk memastikan stabilitas keuangan dan mempertahankan kepercayaan pemangku kepentingan, termasuk investor, dan masyarakat umum [2].

Penilaian risiko kredit dibutuhkan untuk mendeteksi potensi gagal bayar dan menjaga agar portofolio pinjaman tetap sehat. Evaluasi tidak cukup dilakukan sekali di awal karena kondisi debitur dapat berubah seiring waktu, sehingga pemantauan berkelanjutan menjadi langkah yang diperlukan [3]. Sejalan dengan kebutuhan tersebut, kini banyak platform P2P *lending* mulai mengimplementasikan teknik *artificial intelligence* (AI) seperti *machine learning* untuk menganalisis data debitur dalam penilaian risiko kredit, karena dinilai lebih efektif dan akurat dibandingkan dengan metode tradisional seperti regresi logistik [4]. Selain meningkatkan efektivitas dan akurasi prediksi, kemampuan platform dalam menyediakan informasi yang transparan dan dapat dijelaskan juga menjadi faktor penting dalam membangun kepercayaan pengguna, di mana kepercayaan terbukti menjadi faktor utama yang menentukan niat seseorang untuk berinvestasi melalui layanan P2P *lending* [5]. Temuan ini menegaskan bahwa dorongan untuk berinvestasi pada platform P2P *lending* perlu diimbangi dengan transparansi informasi dan mekanisme perlindungan yang jelas [6].

Upaya untuk membangun transparansi dan kepercayaan ini dapat terhambat oleh kompleksitas model AI dan *machine learning* yang sering kali bersifat *black-box* (mekanisme cara kerjanya sulit dipahami). Isu ini menjadi semakin penting mengingat bidang keuangan menuntut tingkat transparansi yang tinggi [7]. Sebagai solusi atas sifat *black-box*, metode seperti *explainable artificial intelligence* (XAI) dikembangkan untuk menghasilkan penjelasan yang dapat dipahami pengguna dengan potensi penerapan pada berbagai tugas keuangan, termasuk salah satunya dalam penilaian risiko kredit. Teknik XAI dinilai dapat membantu pemangku kepentingan memahami faktor yang memengaruhi prediksi model dan meningkatkan akuntabilitas [8].

Namun demikian, penerapan XAI dalam konteks penilaian risiko kredit masih menghadapi permasalahan, terutama terkait kepercayaan dan pemahaman pengguna. Efektivitas XAI dalam meningkatkan kepercayaan masih belum konsisten dan pemahaman tentang dampaknya masih terbatas, sehingga penelitian lebih lanjut mengenai persepsi pengguna diperlukan untuk mendukung implementasi

yang efektif [9]. Persepsi pengguna terhadap XAI terbentuk dari interaksi antara manusia dan sistem, serta dipengaruhi oleh faktor pribadi seperti pengalaman sebelumnya atau sikap terhadap AI [10]. Karena persepsi dan cara berpikir pengguna berperan penting dalam bagaimana mereka memahami, menilai, dan memutuskan untuk terus menggunakan layanan berbasis AI, pengembangan XAI perlu berfokus pada desain yang berpusat pada pengguna untuk meningkatkan penerimaan dan kepercayaan terhadap sistem algoritmik [11].

Penelitian ini diarahkan pada pengembangan model XAI yang mempertimbangkan persepsi pengguna dalam penilaian risiko kredit pada P2P *lending* di Indonesia. Tujuan penelitian ini tidak hanya untuk meningkatkan akurasi, transparansi, dan akuntabilitas prediksi model, tetapi juga untuk memahami bagaimana pengguna menilai dan mempercayai rekomendasi yang dihasilkan. Fokus pada penelitian ini mencakup analisis terhadap persepsi pengguna serta pengembangan model XAI yang akurat dan interpretatif. Pendekatan ini menekankan integrasi antara aspek teknis dan aspek pengguna, sehingga keputusan yang dihasilkan lebih relevan, dapat diterima, dan dipertanggungjawabkan.

2. TINJAUAN PUSTAKA

Peer-to-peer (P2P) *lending* berkembang menjadi komponen penting dalam industri *financial technology* (fintech) karena memungkinkan individu untuk memberikan pinjaman maupun mengajukan pinjaman secara langsung melalui platform digital. Di Indonesia, P2P *lending* menunjukkan pertumbuhan yang sangat pesat sejak kemunculannya pada tahun 2017. Layanan ini mendapat penerimaan luas dari masyarakat karena proses pinjaman yang cepat, mudah diakses, tidak memerlukan agunan dan tanpa melalui perantara lembaga keuangan tradisional [12].

Perkembangan *machine learning* dan *artificial intelligence* (AI) telah meningkatkan kemampuan sektor keuangan dalam melakukan penilaian risiko kredit secara lebih akurat dan efisien, menggantikan metode statistik tradisional yang kurang efektif pada data besar [13]. Hal ini menunjukkan bahwa pendekatan berbasis *machine learning* kini menjadi salah satu kunci dalam memperkuat ketepatan penilaian risiko kredit. Studi menunjukkan bahwa *light gradient boosting machine* (LightGBM) memberikan akurasi prediksi risiko kredit yang lebih tinggi dibandingkan model AI lainnya seperti *decision tree*, *random forest*, ANN, dan CNN, serta model statistik klasik [14].

Penelitian oleh Dastile dan Celik [15] menunjukkan bahwa penerapan XAI terbukti meningkatkan transparansi dalam analisis risiko kredit, di mana metode *shapley additive explanations* (SHAP) memberikan penjelasan yang paling konsisten dan mudah dipahami terhadap kontribusi tiap variabel dibandingkan LIME, Grad-CAM, atau *saliency map*. Menurut Mosca et al. [16], metode SHAP saat ini menjadi kerangka paling populer dalam XAI untuk interpretabilitas.

Penelitian oleh Li et al. [17] membandingkan *logistic regression*, *random forest*, *multilayer perceptron* (*neural network*), *support vector machine*, *naive bayes*, *categorical gradient boosting*, dan LightGBM dalam memprediksi risiko gagal bayar. Hasilnya, LightGBM terbukti memberikan kinerja terbaik dari sisi AUC, akurasi, *F1-score*, serta waktu komputasi. Penelitian tersebut juga menggunakan SHAP dan LIME untuk menjelaskan keputusan model dan menerapkan teknik *random sampling* untuk menangani masalah ketidakseimbangan data, meski begitu masalah ketidakseimbangan tersebut masih tetap ada dan berpotensi memengaruhi kemampuan generalisasi model di kondisi nyata. Menurut Chen et al. [18], ketidakseimbangan data berpotensi memengaruhi efektivitas dan kemampuan generalisasi model, serta menurunkan konsistensi interpretasi yang diberikan oleh metode XAI seperti SHAP maupun LIME. Ketidakseimbangan data menyebabkan model cenderung memfavoritkan kelas mayoritas (nasabah lancar), sehingga kasus gagal bayar yang lebih sedikit bisa terabaikan.

Menurut Xia [19], penggunaan *synthetic minority over-sampling technique* (SMOTE) terbukti efektif mengatasi ketidakseimbangan data pada dataset kredit, memungkinkan model mempelajari pola dari data minoritas sehingga prediksi menjadi lebih akurat dan interpretatif. Studi dari Wongvorachan et al. [20] menunjukkan bahwa kombinasi teknik SMOTE-NC dan *undersampling* adalah yang paling efektif untuk mengatasi ketidakseimbangan ekstrem, *oversampling* biasa cocok untuk ketidakseimbangan sedang, sedangkan *random undersampling* (RUS) murni kurang direkomendasikan karena berisiko kehilangan informasi penting. Menyeimbangkan data minoritas tidak hanya membuat model lebih akurat dan interpretatif, tetapi juga memengaruhi persepsi dan kepercayaan pengguna, yang sering kesulitan memahami cara kerja algoritma. Selain itu, desain antarmuka dan mekanisme sistem dapat turut menentukan tingkat kepercayaan pengguna terhadap teknologi otomatis seperti XAI [21].

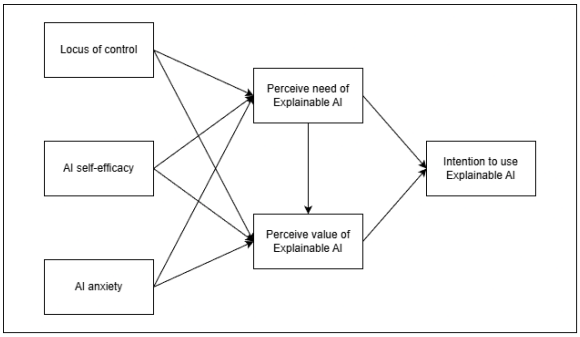
3. METODE PENELITIAN

Penelitian ini menerapkan pendekatan kuantitatif untuk menganalisis faktor dan persepsi pengguna terhadap niat mereka dalam mengadopsi *explainable artificial intelligence* (XAI) untuk penilaian risiko kredit pada *peer-to-peer* (P2P) *lending* di Indonesia, serta pendekatan komputasional untuk merancang dan mengembangkan model XAI yang berorientasi pada pengguna. Tahapan penelitian meliputi identifikasi masalah, studi literatur, pengumpulan dan analisis data kuantitatif melalui kuesioner, pengembangan model, serta evaluasi performa dan interpretasi hasil untuk merumuskan temuan dan rekomendasi pengembangan sistem.

3.1. Analisis Kuantitatif

Untuk menganalisis persepsi dan faktor-faktor yang memengaruhi niat pengguna dalam mengadopsi XAI, penelitian ini menggunakan *technology acceptance model* (TAM) yang telah dimodifikasi sesuai penelitian dari Wang dan Chiou [22] agar sesuai dengan konteks XAI secara umum. Model tersebut kemudian disesuaikan lebih lanjut oleh penulis agar relevan dengan konteks XAI untuk penilaian risiko kredit pada P2P *lending*.

Data kuantitatif dikumpulkan melalui kuesioner yang disebar secara *online*. Populasi target adalah individu yang berpotensi menjadi pengguna (investor) pada platform P2P *lending* di Indonesia dan memiliki pengetahuan dasar tentang *artificial intelligence* (AI). Target jumlah sampel minimum adalah 400 responden dan pengambilan sampel menggunakan teknik *stratified disproportional random sampling*, dibagi menjadi dua strata (pernah dan belum pernah menggunakan P2P *lending*). Adapun model penelitian dapat dilihat pada gambar dibawah ini.



Gambar 1. Model penelitian

Konstruk penelitian berfokus pada beberapa aspek yang memengaruhi niat pengguna dalam menggunakan XAI. Faktor individual meliputi *locus of control*, yaitu keyakinan individu bahwa hasil yang dicapai dipengaruhi oleh tindakan mereka sendiri, AI *self-efficacy*, yang mencerminkan tingkat kepercayaan individu terhadap kemampuan mereka dalam menggunakan AI, dan AI *anxiety*, yaitu kecemasan yang muncul saat menggunakan AI. Selain faktor individual, penelitian ini juga menyoroti persepsi pengguna terhadap teknologi, termasuk *perceived need*, yang menunjukkan seberapa besar kebutuhan pengguna akan XAI, dan *perceived value*, yaitu tingkat penilaian pengguna terhadap manfaat yang diperoleh dari XAI. Semua konstruk tersebut kemudian dianalisis untuk memahami pengaruhnya terhadap niat pengguna dalam mengadopsi XAI.

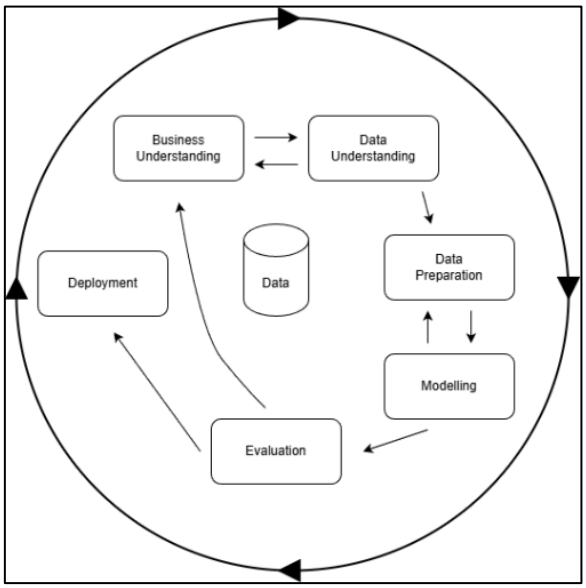
Adapun hipotesis penelitian dirumuskan untuk menguji pengaruh masing-masing faktor individual dan persepsi teknologi terhadap niat pengguna:

H1: *Perceived value* memiliki pengaruh positif terhadap *behavioral intention* XAI.
H2: *Perceived need* memiliki pengaruh positif terhadap *behavioral intention* XAI.
H3: *Perceived need* memiliki pengaruh positif terhadap *perceived value*.
H4: *Internal locus of control* memiliki pengaruh positif terhadap *perceived need* XAI.
H5: *Internal locus of control* memiliki pengaruh positif terhadap *perceived value* XAI.
H6: AI *self-efficacy* memiliki pengaruh positif terhadap *perceived need* XAI.
H7: AI *self-efficacy* memiliki pengaruh positif terhadap *perceived value* XAI.
H8: AI *anxiety* memiliki pengaruh positif terhadap *perceived need* XAI.
H9: AI *anxiety* memiliki pengaruh positif terhadap *perceived value* XAI.

Model penelitian ini diukur menggunakan skala likert 1–5 mulai dari sangat tidak setuju (1) hingga sangat setuju (5). Responden yang tidak memiliki pengetahuan dasar tentang AI nantinya akan disaring dari sampel untuk memastikan validitas temuan, sehingga hasil analisis mencerminkan bahwa responden memiliki pemahaman yang relevan terhadap teknologi yang diteliti. Selanjutnya, data kuantitatif akan dianalisis menggunakan pendekatan *partial least squares-structural equation modeling* (PLS-SEM) melalui perangkat lunak SmartPLS.

3.2. Pengembangan Model

Metode pengembangan model dalam penelitian ini mengikuti kerangka kerja CRISP-DM (*Cross-Industry Standard Process for Data Mining*). Kerangka kerja ini dipilih karena telah terbukti mapan dalam praktik industri dan diakui secara luas sebagai standar untuk pengembangan model data mining [23]. Tahapan CRISP-DM mencakup enam langkah utama: *business understanding*, *data understanding*, *data preparation*, *modeling*, *evaluation*, and *deployment*.



Gambar 2. Kerangka kerja CRISP-DM

Tahap pertama dalam kerangka CRISP-DM diawali dengan *business understanding*, dimana penulis mempelajari pentingnya penggunaan teknik berbasis AI seperti *machine learning* yang transparan dan dapat dijelaskan, serta menganalisis faktor-faktor yang memengaruhi niat pengguna dalam menggunakan XAI untuk penilaian risiko kredit pada platform P2P *lending* di Indonesia. Kemudian, tahap *data understanding* dilakukan dengan mempelajari dataset yang akan digunakan untuk melatih model. Analisis dilakukan untuk menilai kualitas, kelengkapan, dan karakteristik data sebelum melanjutkan ke tahap selanjutnya.

Data preparation dilakukan dengan menangani nilai yang hilang, memilih fitur relevan, mentransformasi data agar siap dianalisis, sekaligus menangani ketidakseimbangan data untuk memastikan model dapat belajar dari distribusi yang seimbang. Setelah data dipersiapkan, tahap *modeling* dilakukan dengan membangun dan melatih model menggunakan dataset yang telah dibersihkan dan dioptimalkan.

Tahap akhir dilakukan melalui *evaluation*, di mana kinerja model dinilai menggunakan metrik yang relevan serta menilai kemampuan model dalam memberikan penjelasan sesuai yang dibutuhkan pengguna. Penelitian ini tidak melanjutkan ke tahap *deployment* karena fokusnya terbatas pada pengembangan dan evaluasi model, tanpa implementasi langsung ke sistem atau platform nyata.

4. HASIL DAN PEMBAHASAN

4.1. Demografi Responden

Hasil penyebaran kuesioner menunjukkan bahwa 403 responden berpartisipasi dalam penelitian ini. Sebanyak 26 responden tidak memiliki pengetahuan mengenai *artificial intelligence* (AI), sehingga total responden yang dapat diikutsertakan dalam proses analisis berjumlah 377 orang. Karakteristik demografis responden secara rinci ditampilkan pada tabel 1.

Tabel 1. Demografi responden

Demografi Responden	Persentase
Jenis kelamin	
Laki-laki	62.8 %
Perempuan	37.2 %
Usia	
≦20	15.0 %
21-30	65.0 %
31-40	15.1 %
41-50	3.2 %
≧51	1.7 %
Pendidikan	
SMA/SMK Sederajat	34.0 %
Diploma / Sarjana (D1-D4, S1)	53.8 %
Magister / S2 ke atas	2.0 %
Lainnya	10.2 %
Pendapatan bulanan	
< Rp 2.000.000	11.9 %
Rp 2.000.001 - Rp 5.000.000	47.9 %
Rp 5.000.001 - Rp 10.000.000	34.2 %
> Rp 10.000.000	6.0 %
Memiliki pegetahuan dasar tentang AI	
Ya	93.5 %
Tidak	6.5 %
Pernah menggunakan P2P lending	
Ya	31.8 %
Tidak	68.2 %

4.2. Uji Realibilitas dan Validitas

Model pengukuran dievaluasi menggunakan empat kriteria, yaitu *indicator realibility*, *internal consistency realibility*, *convergent validity* dan *discriminant validity*. Model pengukuran dinyatakan memenuhi syarat apabila indikator memiliki nilai *outer loading*, *cronbach's alpha* dan rho_A melebihi 0.7, nilai AVE berada di atas 0.5, serta nilai HTMT berada di bawah 0.85 [22].

Indicator reliability dinilai dari *outer loading* tiap indikator. Pada konstruk *locus of control*, beberapa indikator tidak memenuhi ambang batas minimum karena memiliki outer loading rendah, yaitu LC2, LC9, dan LC11, sehingga ketiganya dihapus dari model. Setelah penghapusan, seluruh indikator yang tersisa memiliki *outer loading* kuat (di atas 0.7), sehingga mencerminkan indikator yang valid dan dapat diandalkan.

Tabel 2. Nilai *outer loadings*

Konstruk	Indikator	Outer Loadings
Locus of Control (LC)	LC1	0.854
	LC3	0.859
	LC4	0.872
	LC5	0.839
	LC6	0.860
	LC7	0.832
	LC8	0.839
	LC10	0.852
AI Self-efficacy (ASE)	ASE1	0.882
	ASE2	0.862
	ASE3	0.871
AI Anxiety (AIA)	AIA1	0.758
	AIA2	0.806
	AIA3	0.828
	AIA4	0.763
Perceived Need of XAI (PN)	PN1	0.792
	PN2	0.789
	PN3	0.763
	PN4	0.785
	PN5	0.777
	PN6	0.799
Perceived Value of XAI (PV)	PV1	0.845
	PV2	0.835
	PV3	0.853
Behavioral Intention to use XAI (BI)	BI1	0.931
	BI2	0.911

Internal consistency reability dinilai melalui rho_A dan *cronbach's alpha*. Seluruh konstruk menunjukkan nilai rho_A dan *cronbach's alpha* di atas 0.7, sehingga setiap konstruk memiliki konsistensi internal yang baik dan stabil. *Convergent validity* dievaluasi melalui nilai AVE. Semua konstruk memiliki AVE di atas 0.5, yang berarti varians yang dijelaskan indikator terhadap konstruk lebih besar daripada *error*. Dengan demikian, uji *internal consistency* dan *convergent validity* terpenuhi untuk seluruh konstruk.

Tabel 3. Nilai rho_A, *cronbach's alpha*, dan AVE

Konstruk	rho_A	Cronbach's Alpha	AVE
Locus of Control (LC)	0.965	0.946	0.724
AI Self-efficacy (ASE)	0.847	0.842	0.760
AI Anxiety (AIA)	0.827	0.802	0.623
Perceived Need of XAI (PN)	0.876	0.875	0.615
Perceived Value of XAI (PV)	0.798	0.799	0.713
Behavioral Intention to use XAI (BI)	0.830	0.821	0.848

Discriminant validity dinilai menggunakan nilai HTMT, dengan batas dibawah 0.85. Seluruh pasangan konstruk pada tabel HTMT berada di bawah ambang batas tersebut, sehingga masing-masing konstruk dinyatakan memiliki keterpisahan yang memadai dan tidak saling tumpang tindih.

Tabel 4. Nilai *heterotrait–monotrait ratio* (HTMT)

	LC	ASE	AIA	PN	PV	BI
LC						
ASE	0.163					
AIA	0.156	0.282				
PN	0.109	0.551	0.494			
PV	0.115	0.628	0.278	0.711		
BI	0.197	0.598	0.431	0.425	0.535	

4.3. Evaluasi Model Struktural

Analisis model struktural dilakukan menggunakan teknik *bootstrapping* dengan 5.000 *resamples* untuk menilai signifikansi hubungan antarvariabel dalam penelitian. Evaluasi model dilakukan dengan menggunakan lima indikator utama, yaitu *path coefficient* (β), *t-values*, *p-values*, *f-square*, *variance inflation factor* (VIF), serta *coefficients of determination* (R^2).

Path coefficients (β) menunjukkan seberapa kuat arah dan signifikannya hubungan antara variabel independen dan dependen, serta menggambarkan besarnya pengaruh yang terjadi dalam model penelitian. Sedangkan *t-values* dan *p-values* digunakan untuk menentukan signifikansi statistik pada *path coefficient*, di mana ambang batas umum adalah *t-values* > 1.96 untuk tingkat signifikansi *p-values* 0.05, *t-values* > 2.58 untuk *p-values* 0.01, dan *t-values* > 3.29 untuk *p-values* 0.001. Semakin besar *t-values*, maka semakin kuat bukti pengaruhnya. Semakin kecil *p-values*, semakin kuat keyakinan bahwa hubungan itu nyata. Selain itu, *f-square* digunakan untuk menilai seberapa besar kontribusi masing-masing variabel eksogen terhadap variabel endogen, dengan kategori kecil (0.02–0.15), sedang (0.15–0.35), dan besar (>0.35). VIF digunakan untuk mendeteksi potensi kolinearitas antarkonstruk, nilai di bawah 5 menunjukkan tidak adanya masalah kolinearitas dalam model. Sedangkan, R^2 digunakan untuk menilai kemampuan penjelasan model terhadap variabel dependen, dengan nilai mendekati 0.50 menunjukkan tingkat penjelasan sedang dan mendekati 0.75 menunjukkan kemampuan penjelasan yang kuat [24].

Hasil analisis menunjukkan bahwa variabel *behavioral intention* (BI) memiliki nilai R^2 sebesar 0.206, yang berarti model mampu menjelaskan 20.6% variasi BI. Dari dua prediktor, *perceived value* ($\beta = 0.336$, $t = 3.781$, $p = 0.000$, $f^2 = 0.092$) terbukti berpengaruh signifikan, sedangkan *perceived need* ($\beta = 0.165$, $t = 1.838$, $p = 0.066$, $f^2 = 0.022$) tidak menunjukkan pengaruh signifikan terhadap BI. Dengan demikian, hanya PV yang menjadi penentu utama BI dalam model.

Untuk variabel *perceived need* (PN), nilai R^2 sebesar 0.331 menunjukkan bahwa model mampu menjelaskan 33.1% variasi PN. Hasil pengujian memperlihatkan bahwa AI *self-efficacy* ($\beta = 0.395$, $t = 6.139$, $p = 0.000$, $f^2 = 0.216$) dan AI *anxiety* ($\beta = 0.332$, $t = 7.207$, $p = 0.000$, $f^2 = 0.153$) berpengaruh signifikan, masing-masing dengan ukuran efek sedang. Sebaliknya, *locus of control* tidak signifikan ($\beta = 0.004$, $t = 0.085$, $p = 0.932$, $f^2 = 0.000$). Dengan demikian, PN terutama dipengaruhi oleh ASE dan AIA.

Pada variabel *perceived value* (PV), nilai R^2 sebesar 0.427 menunjukkan bahwa model menjelaskan 42.7% variasi PV. *Perceived need* memberikan pengaruh terbesar ($\beta = 0.469$, $t = 6.586$, $p = 0.000$, $f^2 = 0.257$), diikuti oleh AI *self-efficacy* ($\beta = 0.303$, $t = 5.807$, $p = 0.000$, $f^2 = 0.122$). Sedangkan *locus of control* dan AI *anxiety* tidak memberikan pengaruh signifikan terhadap PV ($p > 0.05$). Dari hasil ini, PV terutama ditentukan oleh PN dan ASE.

Secara keseluruhan, temuan menunjukkan bahwa ASE dan AIA berperan penting dalam meningkatkan PN, sedangkan PN dan ASE berperan dalam meningkatkan PV. Terakhir, PV merupakan satu-satunya faktor signifikan yang meningkatkan BI, menjadikannya konstruk kunci dalam pembentukan niat perilaku penggunaan XAI pada model ini. Adapun rangkuman lengkap hasil pengujian dapat dilihat pada tabel berikut.

Tabel 5. Hasil evaluasi model struktural

Dependant Variable	Independent variable	Path coefficient	t-value	p-value	f-square	VIF	R2
BI	PN	0.165	1.838	0.066	0.022	1.550	0.206
	PV	0.336	3.781	0.000	0.092	1.550	
PN	LC	0.004	0.085	0.932	0.000	1.033	0.331
	ASE	0.395	6.139	0.000	0.216	1.080	
	AIA	0.332	7.207	0.000	0.153	1.074	
PV	PN	0.469	6.586	0.000	0.257	1.495	0.427
	LC	0.015	0.316	0.752	0.000	1.033	
	ASE	0.303	5.807	0.000	0.122	1.313	
	AIA	-0.044	0.753	0.452	0.003	1.238	

4.4. Pemahaman Bisnis

Hasil analisis kuantitatif menunjukkan bahwa persepsi nilai (*perceived value*) menjadi penentu utama niat adopsi (*behavioral intention*) XAI. Persepsi nilai dipengaruhi oleh persepsi kebutuhan (*perceived need*) serta keyakinan diri pengguna dalam menggunakan AI (AI *self-efficacy*). Kecemasan terhadap AI (AI *anxiety*) juga berperan, namun hanya secara tidak langsung melalui peningkatan persepsi kebutuhan, dan efek akhirnya terhadap niat adopsi relatif kecil.

Hal ini sejalan dengan penelitian dari Li et al. [25], yang menyatakan bahwa pengguna dengan *self-efficacy* rendah cenderung tidak percaya diri, mudah bingung, kurang memahami alasan model, dan pada akhirnya membuat keputusan suboptimal. Sedangkan *anxiety* terhadap AI membuat pengguna takut salah, cenderung menghindari eksplorasi, dan menurunkan kemampuan memahami output model.

Menurut Wanner et al. [26], transparansi dalam XAI dapat menurunkan kecemasan pengguna, dimana untuk *end-user*, penjelasan dari *local explanation* lebih efektif dalam membantu memahami keputusan individual dan membangun kepercayaan. Sedangkan *global explanation* justru dapat membebani *end-user* dan meningkatkan kebingungan. Selain itu penelitian Wienrich et al. [27] menunjukkan bahwa perbedaan individual (termasuk *self-efficacy*) memengaruhi bagaimana pengguna merespons penjelasan dari AI. Orang dengan *self-efficacy* tinggi lebih mudah memahami dan melihat nilai dari AI sehingga mereka cenderung menyukai penjelasan yang kompleks, sementara pengguna dengan *self-efficacy* rendah mungkin membutuhkan penjelasan yang lebih sederhana. Hasil dari keseluruhan temuan tersebut pada akhirnya memberikan gambaran dan pemahaman awal mengenai faktor psikologis yang harus dipertimbangkan dalam merancang sistem XAI yang lebih dapat diterima oleh pengguna.

4.5. Pemahaman Data

Dataset yang digunakan berasal dari Bondora, sebuah platform *peer-to-peer* (P2P) *lending* yang menyediakan data publik terkait kinerja pinjaman dan risiko kredit. Data ini mencakup 81.556 baris dan 31 kolom dari tahun 2023. Pemahaman data awal dilakukan melalui *exploratory data analysis* (EDA), di mana penulis mempelajari signifikansi setiap kolom.

Selain itu, penulis juga memeriksa keberadaan data yang hilang dan mengidentifikasi kolom yang tidak relevan atau berpotensi menyebabkan data *leakage*. Pada tahap ini, penulis juga menetapkan kolom *is_default* sebagai target, yaitu kolom yang berisi status *true* (*default)* dan *false* (*non-default)*. *Default* merujuk pada kondisi ketika peminjam gagal memenuhi kewajiban pembayaran, sedangkan *non-default* menunjukkan bahwa pinjaman berjalan lancar tanpa tunggakan. Hasil EDA menunjukkan bahwa distribusi target tidak seimbang (78% *non-default* dan 22% *default*), sehingga diperlukan penanganan untuk menyeimbangkan data sebelum digunakan untuk melatih model.

4.6. Persiapan Data

Pada tahap persiapan data, penulis menangani *missing values* dengan menghapus kolom yang tidak layak digunakan serta memastikan kolom terpilih bebas nilai kosong. *Feature selection* dilakukan dengan mempertahankan hanya fitur yang paling relevan dan bebas data *leakage* yang akan dipakai. Proses *encoding* dilakukan secara minimal, yaitu mengubah variabel target *is_default* dari *true*/*false* menjadi 0/1. Dataset kemudian dibagi menjadi data latih (80%) dan data uji (20%). Karena data tidak seimbang, penulis menerapkan *SMOTE* pada data latih untuk menyeimbangkan kelas *default* dan *non-default* agar model dapat belajar secara lebih optimal. Adapun setelah seluruh proses persiapan data dilakukan, kolom final yang digunakan dalam penelitian ditampilkan pada tabel berikut.

Tabel 6. Fitur yang digunakan untuk melatih model

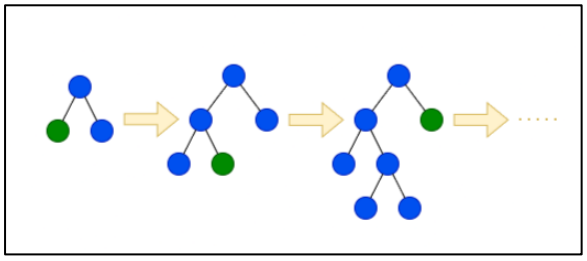
Nama Fitur	Deskripsi
issued_amount	Jumlah pinjaman yang diberikan
initial_interest_rate	Suku bunga awal pinjaman
nr_of_payments	Jumlah cicilan yang sudah dibayar sampai saat ini
principal_balance	Sisa pokok yang belum dibayar
principal_paid_total	Total pokok yang sudah dibayar
extra_interest_paid_total	Total bunga tambahan yang sudah dibayar
late_fee_paid_total	Total denda keterlambatan yang sudah dibayar
initial_loan_duration	Lama pinjaman awal (bulan)
combined_income	Pendapatan gabungan dari data bank dan registri

4.7. Pengembangan dan Pelatihan Model

Sebelum melatih model, penulis melakukan pencarian *hyperparameter* terbaik menggunakan Optuna untuk memperoleh konfigurasi LightGBM yang optimal berdasarkan kombinasi parameter paling efektif untuk dataset. Optuna digunakan untuk LightGBM agar proses pencarian *hyperparameter* menjadi otomatis, efisien, dan lebih akurat dibanding *tuning* manual, sehingga meningkatkan performa model secara signifikan [28]. Proses *tuning* dilakukan pada ruang parameter LightGBM yang mencakup pengaturan seperti *learning rate*, jumlah daun (*num leaves*), kedalaman maksimum (*max depth*), *feature fraction*, *bagging fraction*, *frequency bagging*, jumlah minimum data dalam daun (*min data in leaf*), serta regularisasi lambda L1 dan lambda L2.

Model dengan parameter terbaik kemudian dilatih menggunakan algoritma LightGBM pada data latih yang telah diseimbangkan. Untuk memberikan gambaran mengenai cara kerja algoritma tersebut, berikut ilustrasi arsitektur dan formulasi yang menunjukkan mekanisme *gradient boosting decision trees* pada LightGBM. Setiap pohon dibangun secara bertahap untuk mengoreksi kesalahan dari pohon sebelumnya, sehingga model mampu meningkatkan akurasi prediksi secara progresif:

$$F_m(x) = \sigma\left(\sum_{i=1}^{m} Y_i h_i(x)\right)$$



Gambar 3. Mekanisme *gradient boosting decision trees* pada LightGBM

4.8. Evaluasi Model

Hasil pelatihan model selanjutnya dievaluasi menggunakan data uji untuk mengukur kemampuan prediksi secara objektif. Beberapa metrik seperti AUC, *confusion matrix*, *accuracy*, *precision*, *recall*, dan *F1-score* digunakan untuk menilai kualitas kinerja model dari berbagai aspek.

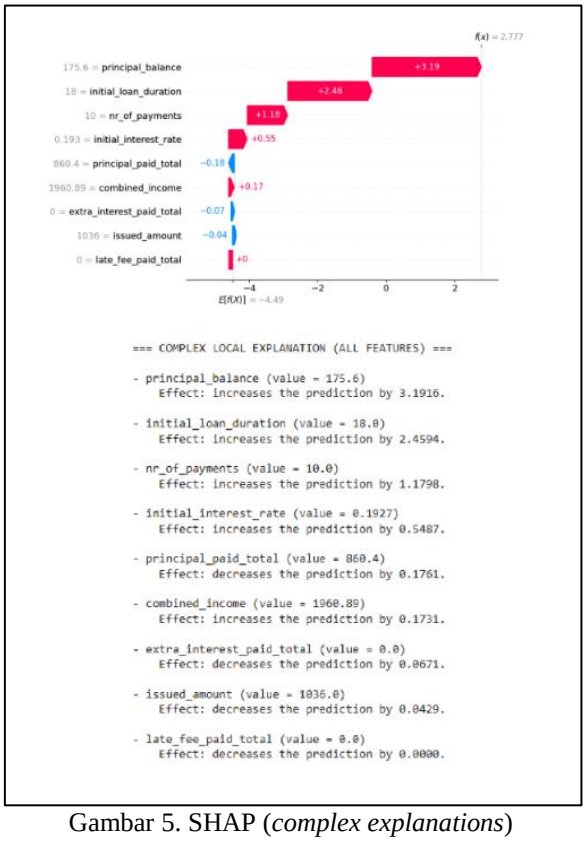
Hasil evaluasi menunjukkan bahwa model LightGBM memiliki skor AUC 0.935. *Confusion matrix* memperlihatkan bahwa sebagian besar kelas *non-default* berhasil diklasifikasikan dengan benar, dan model juga mampu mengenali banyak kasus *default* meskipun masih terdapat 944 *false negative*.

Pada *classification report*, model memperoleh *weighted average* untuk *precision* 0.87, *recall* 0.87, *f1-score* 0.87, dan *accuracy* 0.87. Visualisasi hasil evaluasi dapat dilihat pada gambar di bawah.

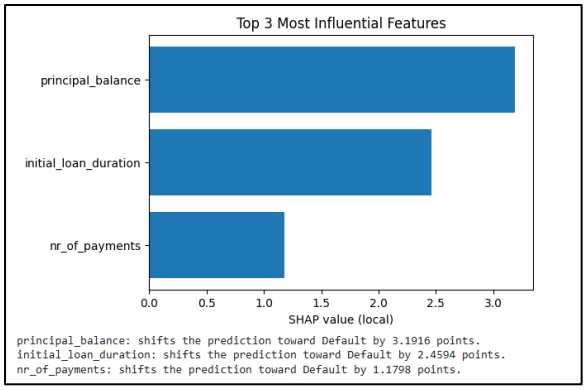
```
ROC-AUC: 0.9350304883555705
-----
Confusion Matrix:
 [[11608  1168]
 [  944  2592]]
-----

Classification Report:
              precision    recall  f1-score   support

           0       0.92      0.91      0.92     12776
           1       0.69      0.73      0.71      3536

    accuracy                           0.87     16312
   macro avg       0.81      0.82      0.81     16312
weighted avg       0.87      0.87      0.87     16312
```

Gambar 4. Hasil evaluasi model

Untuk interpretasi model penulis menggunakan metode SHAP. Menurut penelitian dari Gramegna dan Giudici [29], dalam konteks penilaian risiko kredit, SHAP lebih disarankan dibanding LIME karena mampu memberikan penjelasan yang lebih jelas dan konsisten tentang pengaruh setiap fitur terhadap prediksi *default*, sehingga memudahkan pemahaman akan dinamika risiko kredit. Penjelasan ini disajikan dalam dua tipe: *complex explanation* untuk penjelasan detail yang lebih mendalam dan *simple explanation* untuk penjelasan yang lebih ringkas dan sederhana.



Gambar 5. SHAP (*complex explanations*)



Gambar 6. SHAP (*simple explanations*)

5. KESIMPULAN DAN SARAN

Hasil penelitian menunjukkan bahwa penerimaan pengguna terhadap XAI untuk penilaian risiko kredit pada konteks P2P *lending* di Indonesia menuntut bentuk penjelasan yang selaras dengan kebutuhan dan karakteristik psikologis pengguna, khususnya terkait kemampuan mereka dalam memahami dan menggunakan AI. Selain itu, penyajian penjelasan lokal juga diperlukan untuk mengurangi kecemasan dan meningkatkan pemahaman terhadap hasil yang diberikan oleh AI. Model yang dikembangkan menggunakan LightGBM dan dijelaskan melalui SHAP menghasilkan akurasi dan transparansi dalam penilaian risiko kredit, serta mampu memenuhi pola persepsi pengguna yang teridentifikasi dalam analisis kuantitatif. Untuk penelitian selanjutnya, disarankan agar model dapat diuji melalui skenario penerapan nyata baik melalui *deployment* maupun pengujian terkontrol dengan menggunakan dataset lokal, sehingga prediksi dan interpretasinya dapat lebih relevan dan siap digunakan dalam konteks operasional nyata.

DAFTAR PUSTAKA

[1] M. Annas and M. A. Ansori, "Problems in Determining Interest in Peer-to-Peer Lending in Indonesia," *Jurnal Media Hukum*, vol. 28, no. 1, pp. 102–116, Jun. 2021, doi: 10.18196/jmh.v28i1.10022.

[2] S. Anuwak, K. Sintanakul, and C. Sanrach, "Efficiency Enhancement with Rule-Based Method for Credit Classification," *ASEAN Journal of Scientific and Technological Reports*, vol. 25, no. 2, pp. 10–18, Apr. 2022, doi: 10.55164/ajstr.v25i2.244172.

[3] M. R. Machado, D. T. Chen, and J. R. Osterrieder, "An analytical approach to credit risk assessment using machine learning models," *Decision Analytics Journal*, vol. 16, Sep. 2025, doi: 10.1016/j.dajour.2025.100605.

[4] A.-H. Chang, L.-K. Yang, R.-H. Tsaih, and S.-K. Lin, "Machine learning and artificial neural networks to construct P2P lending credit-scoring model: A case using Lending Club data," *Quantitative Finance and Economics*, vol. 6, no. 2, pp. 303–325, 2022, doi: 10.3934/qfe.2022013.

[5] Mudjahidin, A. A. Hidayat, and A. P. Aristio, "Conceptual model of use behavior for peer-to-peer lending in Indonesia," in *Procedia Computer Science*, Elsevier B.V., 2021, pp. 215–222. doi: 10.1016/j.procs.2021.12.134.

[6] R. Rusdiyanto, R. Sahid, D. Hardyansah, D. S. Darmawan, R. Negara, and K. Khan, "Legal Perspective of Investment Risk Mitigation on Peer-to-peer Lending Platforms," 2023.

[7] J. Černevičienė and A. Kabašinskas, "Explainable artificial intelligence (XAI) in finance: a systematic literature review," *Artif Intell Rev*, vol. 57, no. 8, Aug. 2024, doi: 10.1007/s10462-024-10854-8.

[8] F. S. Khan, S. S. Mazhar, K. Mazhar, D. A. AlSaleh, and A. Mazhar, "Model-agnostic explainable artificial intelligence methods in finance: a systematic review, recent developments, limitations, challenges and future directions," *Artif Intell Rev*, vol. 58, no. 8, Aug. 2025, doi: 10.1007/s10462-025-11215-9.

[9] P. Hamm, M. Klesel, P. Coberger, and H. F. Wittmann, "Explanation matters: An experimental study on explainable AI," *Electronic Markets*, vol. 33, no. 1, Dec. 2023, doi: 10.1007/s12525-023-00640-9.

[10] T. Ha, Y. J. Sah, Y. Park, and S. Lee, "Examining the effects of power status of an explainable artificial intelligence system on users' perceptions," *Behaviour and Information Technology*, vol. 41, no. 5, pp. 946–958, 2022, doi: 10.1080/0144929X.2020.1846789.

[11] D. Shin, "The effects of explainability and causability on perception, trust, and acceptance: Implications for explainable AI," *International Journal of Human Computer Studies*, vol. 146, Feb. 2021, doi: 10.1016/j.ijhcs.2020.102551.

[12] M. Y. Edward, E. N. Fuad, H. Ismanto, A. D. R. Atahau, and Robiyanto, "Success factors for peer-to-peer lending for SMEs: Evidence from Indonesia," *Investment Management and Financial Innovations*, vol. 20, no. 2, pp. 16–25, 2023, doi: 10.21511/imfi.20(2).2023.02.

[13] S. Shi, R. Tse, W. Luo, S. D'Addona, and G. Pau, "Machine learning-driven credit risk: a systemic review," Sep. 01, 2022, *Springer Science and Business Media Deutschland GmbH*. doi: 10.1007/s00521-022-07472-2.

[14] P. C. Ko, P. C. Lin, H. T. Do, and Y. F. Huang, "P2P Lending Default Prediction Based on AI and Statistical Models," *Entropy*, vol. 24, no. 6, Jun. 2022, doi: 10.3390/e24060801.

[15] X. Dastile and T. Celik, "Making Deep Learning-Based Predictions for Credit Scoring Explainable," *IEEE Access*, vol. 9, pp. 50426–50440, 2021, doi: 10.1109/ACCESS.2021.3068854.

[16] E. Mosca, F. Szigeti, S. Tragianni, D. Gallagher, and G. Groh, "SHAP-Based Explanation Methods: A Review for NLP Interpretability,"

2022. [Online]. Available: https://github.com/slundberg/shap

[17] L. H. Li, A. K. Sharma, and S. T. Cheng, "Explainable AI based LightGBM prediction model to predict default borrower in social lending platform," *Intelligent Systems with Applications*, vol. 26, Jun. 2025, doi: 10.1016/j.iswa.2025.200514.

[18] Y. Chen, R. Calabrese, and B. Martin-Barragan, "Interpretable machine learning for imbalanced credit scoring datasets," *Eur J Oper Res*, vol. 312, no. 1, pp. 357–372, Jan. 2024, doi: 10.1016/j.ejor.2023.06.036.

[19] H. Xia, W. An, and Zuopeng (Justin) Zhang, "Credit Risk Models for Financial Fraud Detection," *Journal of Database Management*, vol. 34, no. 1, pp. 1–20, Apr. 2023, doi: 10.4018/jdm.321739.

[20] T. Wongvorachan, S. He, and O. Bulut, "A Comparison of Undersampling, Oversampling, and SMOTE Methods for Dealing with Imbalanced Classification in Educational Data Mining," *Information (Switzerland)*, vol. 14, no. 1, Jan. 2023, doi: 10.3390/info14010054.

[21] D. Ribes *et al.*, "Trust Indicators and Explainable AI: A Study on User Perceptions," in *Lecture Notes in Computer Science (including subseries Lecture Notes in Artificial Intelligence and Lecture Notes in Bioinformatics)*, Springer Science and Business Media Deutschland GmbH, 2021, pp. 662–671. doi: 10.1007/978-3-030-85616-8_39.

[22] Y. M. Wang and C. C. Chiou, "Exploring Factors Affecting User Intention to Accept Explainable Artificial Intelligence," *Computer Science and Information Systems*, vol. 22, no. 3, pp. 1081–1104, Jun. 2025, doi: 10.2298/CSIS241018041W.

[23] U. Kannengiesser and J. S. Gero, "MODELLING THE DESIGN OF MODELS: AN EXAMPLE USING CRISP-DM," in *Proceedings of the Design Society*, Cambridge University Press, 2023, pp. 2705–2714. doi: 10.1017/pds.2023.271.

[24] C. H. Huang, "Using pls-sem model to explore the influencing factors of learning satisfaction in blended learning," *Educ Sci (Basel)*, vol. 11, no. 5, May 2021, doi: 10.3390/educsci11050249.

[25] X. Li, Q. Zong, and M. Cheng, "The Impact of Medical Explainable Artificial Intelligence on Nurses' Innovation Behaviour: A Structural Equation Modelling Approach," *J Nurs Manag*, vol. 2024, 2024, doi: 10.1155/2024/8885760.

[26] J. Wanner, L. V. Herm, K. Heinrich, and C. Janiesch, "The effect of transparency and trust on intelligent system acceptance: Evidence from a user-based study," *Electronic Markets*, vol. 32, no. 4, pp. 2079–2102, Dec. 2022, doi: 10.1007/s12525-022-00593-5.

[27] C. Wienrich, A. Carolus, D. Roth-Isigkeit, and A. Hotho, "Inhibitors and Enablers to Explainable AI Success: A Systematic Examination of Explanation Complexity and Individual Characteristics," *Multimodal Technologies and Interaction*, vol. 6, no. 12, Dec. 2022, doi: 10.3390/mti6120106.

[28] K. Mo *et al.*, "Fresh Meat Classification Using Laser-Induced Breakdown Spectroscopy Assisted by LightGBM and Optuna," *Foods*, vol. 13, no. 13, Jul. 2024, doi: 10.3390/foods13132028.

[29] A. Gramegna and P. Giudici, "SHAP and LIME: An Evaluation of Discriminative Power in Credit Risk," *Front Artif Intell*, vol. 4, 2021, doi: 10.3389/frai.2021.752558