

## MODEL KELAYAKAN KREDIT PADA P2P LENDING MENGGUNAKAN DATA SOSIAL MEDIA

**Raynald Damara Eriyanto<sup>1)</sup>, Alvin Ramaquita<sup>2)</sup>, Andry Alamsyah<sup>3)</sup>**

<sup>1,2,3</sup>Fakultas Ekonomi dan Bisnis, Universitas Telkom

Email :<sup>1</sup>raynalddamara98@gmail.com, <sup>2</sup>alvinrama@telkomuniversity.ac.id, <sup>3</sup>andrya@telkomuniversity.ac.id

### Abstrak

Lembaga keuangan cenderung lebih suka memiliki peminjam dengan sumber riwayat kredit yang jelas. Berbeda dengan *P2P Lending* yang tidak memiliki data secara lengkap, *P2P Lending* menggunakan data alternatif atau pengganti untuk membuat model penilaian kredit. Sosial media menjadi salah satu opsi data pengganti tersebut tidak dapat menerjemahkan secara langsung konten terkait permasalahan kredit. Oleh karena itu diperlukan adanya pendekatan dan pemahaman data sosial media. Dengan menggunakan data *user demographic attributes* atau fitur demografi sosial media, penelitian ini bertujuan untuk membuat model penilaian kredit menggunakan data pengganti berupa data sosial media bagi *P2P Lending*. Penelitian ini menggunakan proses *data mining* yang memanfaatkan algoritma *decision tree* dan *random forest*. Hasil penelitian menunjukkan model penilaian kredit dengan algoritma *random forest* menghasilkan *accuracy* tertinggi sebesar 81.25% dan dipilih menjadi model terbaik. Dari hasil tersebut, *P2P Lending* dapat memiliki kesempatan untuk membuka segmentasi konsumen baru yang tidak memiliki Riwayat keuangan atau kredit secara lengkap.

**Kata Kunci :** Kredit, Penilaian Kredit, *P2P Lending*, *Data Mining*.

### Abstract

*Financial institutions tend to prefer to have borrowers with a clear source of credit history. Unlike P2P Lending which does not have complete data, P2P Lending uses alternative or substitute data to create a credit scoring model. Social media as one of the alternative data options cannot directly translate content related credit problems. Therefore, it is necessary to have an approach and understanding of social media data. By using data user demographic attributes or social media demographic features, this study aims to create a credit scoring model using substitute data in the form of social media data for P2P Lending. This study uses a data mining process that utilizes a decision tree algorithm and random forest. The results showed that the credit scoring model with the random forest algorithm produced the highest accuracy of 81.25% and was chosen to be the best model. From these results, P2P Lending can have the opportunity to open a new segment of consumers who do not have a complete financial or credit history.*

**Keywords:** Credit, Credit Scoring, *P2P Lending*, *Data Mining*.

### 1. PENDAHULUAN

Lembaga keuangan melakukan penilaian kredit berpedoman pada analisis 5C yang meliputi *character*, *capacity*, *capital*, *condition*, dan *collateral* dengan prinsip 5P yaitu, *personality*, *purpose*, *prospect*, *payment*, dan *party* (Fahmi, 2014). Lembaga keuangan cenderung lebih suka memiliki peminjam dengan sumber data riwayat kredit yang jelas dan data lain untuk memperkuat skor kredit peminjam. Berbeda dengan *P2P Lending* yang tidak memiliki data peminjam secara lengkap, *P2P*

*Lending* mengandalkan data lain atau disebut data alternatif untuk membuat penilaian kredit tersendiri (DailySocial.id, 2019).

*P2P (peer to peer) Lending* merupakan sebuah layanan keuangan secara online yang mempertemukan peminjam dan pencari dana (koinworks.com, 2016). Di Indonesia, *Fintech Lending* mengalami pertumbuhan yang sangat cepat, sejak tanggal 31 Januari 2020, Otoritas Jasa Keuangan (OJK) mencatat sebanyak 164 perusahaan terdaftar dengan 25 perusahaan diantaranya telah

memiliki izin resmi OJK. OJK juga mencatat total akumulasi peminjam *Fintech Lending* di Indonesia per Januari 2020 sebesar 20.497.167 orang atau naik 292,83% dari periode tahun sebelumnya dengan persentase terbesar oleh laki-laki sebesar 51,21% dan rentang usia peminjam 19-34 tahun sebesar 70,33% (Otoritas Jasa Keuangan, 2020).

Tingginya jumlah peminjam *Fintech Lending* tidak lepas dari beberapa keunggulan yang dimilikinya seperti kemudahan dalam pengajuan dana pinjaman, keamanan yang terjamin jika terdaftar atau mendapat izin usaha dari OJK, transparansi dalam memberikan laporan pendanaan, dan dapat menjangkau masyarakat *unbanked* (Amartha, 2019). Akan tetapi *P2P Lending* juga memiliki permasalahan, menurut pendapat Guo *et al.* (2016) yang menyatakan *P2P Lending* memiliki tantangan dalam menggunakan metode penilaian kredit yang hampir sama dengan metode penilaian kredit secara tradisional yaitu, *data coverage* (cakupan data), *data timeliness* (jangka waktu data), dan *data availability* (ketersediaan data). Berkaitan dengan *data availability*, perusahaan *P2P Lending* masih mengalami kesulitan untuk mendapatkan akses catatan kredit atau data keuangan konsumen dibandingkan lembaga keuangan lainnya yang menyebabkan adanya kekurangan data untuk menentukan penilaian kredit (Guo *et al.*, 2016).

Ketersediaan data yang tidak mencukupi memberikan tantangan kepada perusahaan *P2P Lending* untuk memanfaatkan data alternatif sebagai pengganti atau pelengkap indikator pemodelan penilaian kredit. Data sosial media yang dihasilkan peminjam akhirnya menjadi alternatif data yang penting bagi penilaian kelayakan kredit perusahaan *P2P Lending*. Akan tetapi, data sosial media yang heterogen seperti unggahan artikel atau konten, koneksi pertemanan, dan sejenisnya tidak mengandung konten terkait kredit secara eksplisit dalam banyak kasus (Hua *et al.*, 2015). Sedikitnya relevansi antara data sosial media dengan konten terkait penilaian kredit memerlukan adanya penyesuaian dan pengolahan data yang tepat. Pengolahan data sosial perlu dilakukan dengan memanfaatkan metode statistika secara maksimal, sebagian literatur penilaian kredit sebelumnya merumuskannya sebagai masalah klasifikasi biner (Thomas *et al.*, 2002; dalam Guo *et al.*, 2016). Dengan memanfaatkan data sosial media berupa data *user demographic attributes* yang nantinya akan diubah menjadi fitur demografi, penelitian ini memiliki tujuan untuk membuat model penilaian kelayakan kredit menggunakan fitur demografi pada sosial media.

## 2. KAJIAN PUSTAKA

### 2.1 Kredit

Berdasarkan Pedoman Akuntansi Perbankan Indonesia (PAPI) tahun 2001 dalam Fahmi (2014) kredit didefinisikan sebagai penyediaan uang atau tagihan yang dapat dipersamakan dengan itu, berdasarkan persetujuan atau kesepakatan pinjam meminjam antara bank dan pihak lain yang mewajibkan pihak peminjam (debitur) untuk melunasi utangnya setelah jangka waktu tertentu dengan jumlah bunga, imbalan atau pembagian hasil keuntungan. Kredit juga dapat didefinisikan sebagai sejumlah dana yang dipinjamkan antara 2 orang atau lebih sesuai kesepakatan dan ketentuan yang telah disepakati bersama.

### 2.2 Credit Scoring (Penilaian Kredit)

Menurut Ikatan Bankir Indonesia (2015) setiap Lembaga keuangan memiliki cara dan proses penilaian kredit yang berbeda-beda yang dapat disesuaikan sesuai kebutuhan dan kepentingan masing-masing lembaga tersebut. Secara umum setiap lembaga keuangan berpedoman pada prinsip kredit 5C, yaitu *capital*, *collateral*, *condition*, *character*, dan *capacity* dalam menentukan penilaian kreditnya.

### 2.3 Decision Tree

*Decision tree* merupakan algoritma yang popular. Algoritma ini menggunakan struktur pohon atau hierarki dalam melakukan prediksi suatu model. Algoritma ini merupakan struktur berurutan yang terdiri dari *nodes* (simpul) dan *branches* (cabang/ujung yang terarah) yang menghubungkan simpul. *Decision tree* memiliki kelebihan dalam membuat suatu model untuk pertama kali karena metode ini memadukan pemodelan dan eksplorasi data yang ada. Algoritma ini juga memiliki kekurangan dalam mengerjakan model dengan kriteria atau kelas yang banyak karena bisa menyebabkan *overlap* (Suyanto, 2019).

### 2.4 Random Forest

*Random forest* merupakan algoritma yang berasal dari pengembangan *decision tree*. Algoritma *random forest* dapat dikatakan sebagai gabungan banyak algoritma *decision tree* yang bertujuan untuk mendapatkan hasil prediksi yang lebih maksimal atau meminimalisir kesalahan pada *decision tree* (Breiman, 2001). Algoritma ini merupakan *ensemble learning* yang bekerja meminimalisir kesalahan dan mereduksi varian dengan membagi rata suatu kumpulan data agar memiliki tingkat korelasi yang rendah sehingga dimaksudkan untuk meningkatkan proses *learning* dalam suatu model (Wibowo *et al.*, 2016).

## 2.5 Fitur Demografi

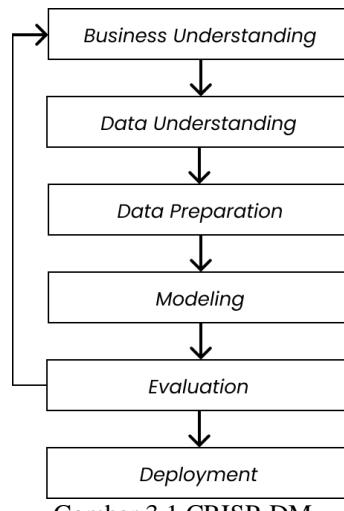
Fitur demografi atau data *user demographic attributes* merupakan kumpulan data terkait pribadi yang ditambahkan oleh seseorang di akun sosial media yang mencerminkan karakteristik demografinya. Menurut Guo *et al* (2016) data tersebut dapat berupa jenis kelamin, pekerjaan, riwayat pendidikan, tempat tinggal dan lainnya. Beberapa data sosial media secara tidak langsung dapat mencerminkan prinsip pokok terkait kredit, hal tersebut didapat dengan memahami data lebih lanjut. Umur dapat mencerminkan stabilitas ekonomi seseorang yang dapat mewakili prinsip *capacity*. Begitu pula riwayat kerja yang mencerminkan pengalaman seseorang dalam karirnya yang dapat mewakili prinsip *capacity*. Riwayat pendidikan yang dapat mencerminkan Pendidikan yang baik yang mewakili prinsip *capacity, character, dan conditions*.

Tabel 2.1 Fitur Demografi

Indikator	Prinsip Kredit
Umur (age)	<i>Capacity</i>
Gaji (salary)	<i>Capacity</i>
Jenis kelamin (gender)	<i>Capacity</i>
Pendidikan (education)	<i>Capacity, character, dan conditions.</i>

## 3. METODE PENELITIAN

Penelitian ini memanfaatkan metode CRISP-DM atau *Cross-Industry Standard Process for Data Mining* yang standarisasi *data mining* yang disusun oleh tiga pengagas *data mining market* yaitu Daimler Chrysler (Daimler-Benz), SPSS (ISL), NCR (Shearer, 2000; dalam Fadillah, 2015).



Gambar 3.1 CRISP-DM

### a. Business Understanding

Tahap ini fokus pada tujuan akhir dilakukannya penelitian. Penelitian ini memiliki tujuan untuk membuat model penilaian kredit menggunakan data sosial media.

### b. Data Understanding

Pada tahap ini, dilakukan pengambilan data dan pemahaman akan data. Data pada penelitian ini menggunakan data sosial media LinkedIn yang nantinya proses pengambilan memanfaatkan jasa web *scraping* Phantombuster. Pada Phantombuster, digunakan 2 fitur yang tersedia, yaitu *LinkedIn Followers Insights* untuk mengambil data koneksi pertemanan, dan *LinkedIn Profile Scraper* untuk mengambil data profil setiap akun yang masuk kedalam jaringan pertemanan, nantinya data ini akan menjadi data *user demographic attributes* atau fitur demografi. Pada tahap ini menghasilkan data 3578 akun LinkedIn yang masih memiliki banyak kekosongan data.

### c. Data Preparation

Pada tahap ini, dilakukan penyaringan, penggabungan dan pemilihan data untuk mendapatkan data terbaik bagi penelitian. Pemilihan data dilakukan dengan memilih data yang memiliki kelengkapan atribut data *company, job title, dan school*. Pada tahap ini menghasilkan 112 akun LinkedIn yang memiliki kelengkapan data yang diinginkan. Dari data tersebut, dilakukan beberapa pendekatan untuk mendapatkan data terkait penilaian kredit berdasarkan fitur demografi.

Tabel 3.1 Ekstraksi Data Atribut

Atribut Data	Data Hasil Ekstraksi
<i>full name</i> dan <i>imgUrl</i>	<i>Gender</i>
<i>company, job title</i>	<i>Age</i> dan <i>school</i>
<i>school</i>	<i>Education</i>
<i>company</i> dan <i>job title</i>	<i>Salary</i>

### d. Modeling

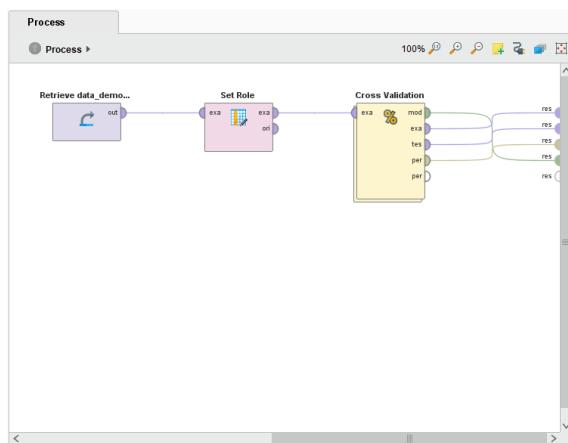
Pada tahap ini, akan dilakukan pemilihan model dan membangun model. Penelitian ini menggunakan *data mining* dengan tujuan untuk melakukan *prediction analysis* berupa metode *classification* (klasifikasi) yang memanfaatkan algoritma *decision tree* dan *random forest*. Pemilihan metode klasifikasi dilakukan karena data yang telah

didapat akan diberikan label kelayakan kredit sesuai *threshold* yang akan ditentukan nantinya.

Tabel 3.2 *Credit Scoring Scorecard*

No	Atribut	Nilai	Keterangan
1	Salary	1	0-5.000.000
		2	5.000.0001-
		3	10.000.000
		4	10.000.001-
		5	20.000.000
			20.000.001-
			40.000.000
			>40.000.000
		*dalam rupiah	
2	Age	1	<19
		2	19-34
		3	35-54
		4	>54
3	Gender	1	Wanita
		2	Pria
4	Education	1	Kluster 5
		2	Kluster 4
		3	Kluster 3
		4	Kluster 2
		5	Kluster 1

Setelah data diberikan skor sesuai dengan table 3.2, selanjutnya setiap data akan diberikan label sesuai *threshold* yang telah ditentukan sebelumnya. Kemudian data akan diolah menggunakan bantuan *software* Rapidminer.



Gambar 3.2 Proses Rapidminer

#### e. Evaluation

Pada tahap ini, hasil akan mempresentasikan sejauh mana sebuah rancangan model telah memenuhi tujuan awal penelitian. Tahap validasi menggunakan *cross validation* pada *software* Rapidminer, *cross validation* merupakan pembanding algoritma pembelajaran dengan membagi data menjadi dua bagian secara statistic untuk menguji validasi terhadap suatu model (Refaeilzadeh *et al.*, 2009). Selanjutnya hasil kinerja model akan dievaluasi oleh *confusion metrics* yang menghasilkan *accuracy*, *precision*, *recall*, dan *f-measure* yang merupakan *performance metrics*.

#### f. Deployment

Pada tahap ini, akan dibuat laporan akhir tentang hasil penelitian yang ada, tahap ini akan menentukan model akhir yang nantinya dapat diimplementasikan sesuai tujuan awal penelitian.

### 4. HASIL DAN PEMBAHASAN

Dalam penelitian ini, data diambil menggunakan bantuan jasa *web scraping* Phantombuster. Data yang diambil berasal dari sosial media LinkedIn yang berjumlah 3578 akun dan nantinya diambil 112 akun setelah melewati tahap *data preparation*. Dari data tersebut nantinya akan diambil wawasan-nya untuk dijadikan fitur demografi.

Tabel 4.1 Fitur Sosial Media

Fitur Sosial Media	Atribut Data
Fitur Demografi	<i>Salary</i>
	<i>Age</i>
	<i>Gender</i>
	<i>Education</i>

*Salary* didapat dengan melakukan pemahaman lebih lanjut pada atribut data *company* dan *job title*. Pada atribut data tersebut dapat diketahui posisi jabatan kerja seseorang dan besaran gaji yang didapat, nilai gaji yang diperoleh dari hasil pendekatan yang dilakukan dengan membandingkan laporan tahunan dari 3 sumber yaitu, laporan tahunan Kelly dan MichaelPage beserta *web* Glasdoor.com.

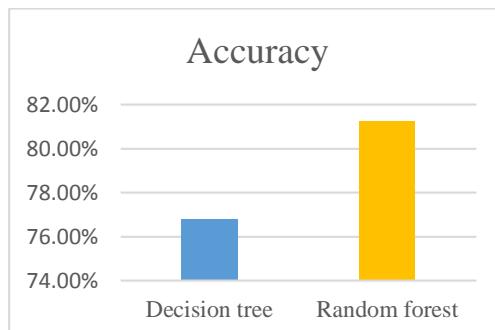
*Age* didapat dengan melakukan pemahaman lebih lanjut pada atribut data *company*, *job title*, and *school*. Pada atribut data tersebut dapat diketahui tahun seseorang menyelesaikan Pendidikan dengan asumsi yang telah ditentukan dan jabatan kerja yang didapat membantu memperkirakan lama kerja seseorang.

*Education* didapat dengan melakukan pemahaman lebih lanjut pada atribut data *school*. Pada atribut tersebut dapat diketahui riwayat pendidikan yang telah ditempuh seseorang. *Gender* didapat dengan melakukan pemahaman lebih lanjut

pada atribut data *full name* dan *imgUrl* yang mana akan mengarahkan pada nama dan gambar profil yang terpasang pada akun sosial media.

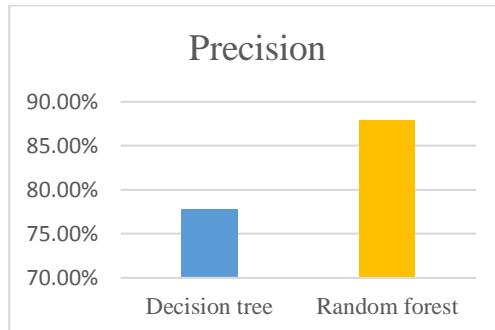
Setelah data disiapkan, maka dibutuh model percobaan penilaian kredit. Penelitian ini terlebih dahulu menentukan *threshold* atau batasan model penilaian kredit dari fitur demografi yang ada. Setelah itu data akan diberikan bobot skor sesuai dengan nilai dari masing-masing atribut data yang ada. Data yang telah diberi bobot skor kemudian diberi label sesuai dengan skor yang diperoleh, jika skor yang diperoleh melebihi *threshold* yang ditentukan maka akan mendapat label *Eligible* atau layak mendapatkan kredit dan begitu sebaliknya. Setelah melakukan pelabelan, maka data diolah menggunakan algoritma *decision tree* dan *random forest* untuk dievaluasi kinerja *performance metrics* masing-masing algoritma.

Setelah data diolah menggunakan *software* RapidMiner, akan dapat diketahui hasil *performance metrics* tiap algoritma yang ada. Algoritma *decision tree* menghasilkan *accuracy* sebesar 76.79% dan algoritma *random forest* menghasilkan *accuracy* sebesar 81.25%.



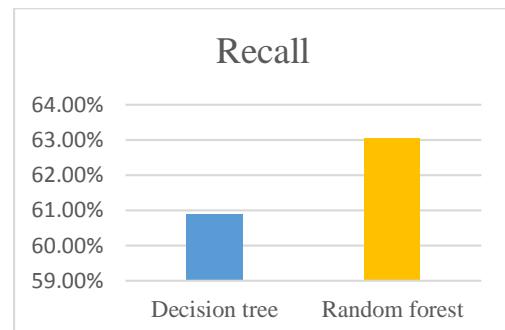
Gambar 4.1 Hasil Accuracy

Algoritma *decision tree* menghasilkan nilai *precision* sebesar 77.78% dan algoritma *random forest* menghasilkan *precision* sebesar 87.88%.



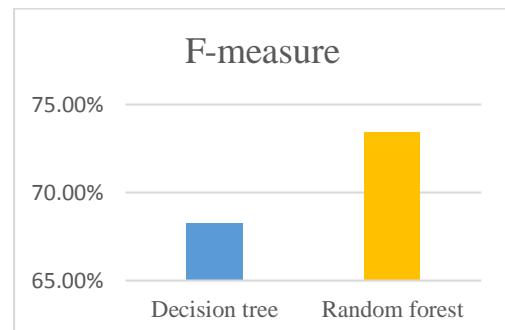
Gambar 4.2 Hasil Precision

Algoritma *decision tree* menghasilkan nilai *recall* sebesar 60.87% dan algoritma *random forest* menghasilkan *recall* sebesar 63.04%.



Gambar 4.3 Hasil Recall

Algoritma *decision tree* menghasilkan nilai *f-measure* sebesar 68.29% dan algoritma *random forest* menghasilkan *f-measure* sebesar 73.42%.



Gambar 4.4 Hasil F-measure

Dapat dilihat hasil dari *confusion metrics* algoritma *decision tree* pada gambar 4.5 bahwa algoritma tersebut menghasilkan *True Positive* sebanyak 58 data, *True Negative* sebanyak 28 data, *False Positive* sebanyak 18 data, dan *False Negative* sebanyak 8 data. Dari hasil tersebut dapat diketahui algoritma *random forest* mampu memprediksi dengan benar sejumlah 86 data dan salah memprediksi sebanyak 26 data.

	true 0	true 1	class precision
pred. 0	58	18	76.32%
pred. 1	8	28	77.78%
class recall	87.88%	60.87%	
class recall			

Gambar 4.5 Confusion metrics decision tree

Dapat dilihat hasil dari *confusion metrics* algoritma *random forest* pada gambar 4.6 bahwa algoritma tersebut menghasilkan *True Positive* sebanyak 62 data, *True Negative* sebanyak 29 data, *False Positive* sebanyak 17 data, dan *False Negative* sebanyak 3 data. Dari hasil tersebut dapat diketahui algoritma *random forest* mampu memprediksi

dengan benar sejumlah 91 data dan salah memprediksi sebanyak 21 data.

	true 0	true 1	class precision
pred 0	62	17	78.48%
pred 1	4	29	87.88%
class recall	93.94%	63.04%	

Gambar 4.6 Confusion metrics random forest

Berdasarkan perbandingan tiap *performance metrics* dari kedua algoritma tersebut, algoritma *random forest* memiliki kinerja paling tinggi dalam menilai kelayakan kredit karena unggul disemua *performnce metrics* yang ada. Oleh karena itu algoritma *random forest* merupakan pilihan terbaik untuk membuat model kelayakan kredit yang menggunakan fitur demografi sosial media.

Tingginya hasil *performance metrics* pada algoritma *random forest* bisa disebabkan karena algoritma tersebut mampu mereduksi varian yang ada dalam data, dengan membagi rata setiap data dan memisahkan data tersebut ke banyak pohon keputusan dapat meminimalisir korelasi antar data sehingga mampu menghasilkan keputusan terbaik dari banyak kemungkinan yang ada. Sedangkan pada algoritma *decision tree* hasil yang didapat tidak setinggi algoritma *random forest* dimana hal tersebut bisa terjadi karena adanya *overlap*, terutama bila kelas dan jumlah data yang digunakan terlalu bervariasi dan berjumlah banyak. Selain itu, bisa disebabkan karena kurang optimalnya desain model yang dibuat.

Dari hasil *performance metrics* kedua algoritma yang digunakan dapat diambil wawasan bahwa fitur demografi sosial media atau data *user demographic attributes* dapat dijadikan data alternatif atau pengganti dalam model penilaian kelayakan kredit pada *P2P Lending*. Hal tersebut juga dapat menjadi peluang bagi *P2P Lending* dalam membuka segmentasi konsumen yang baru, terutama konsumen yang tidak memiliki riwayat keuangan atau kredit yang lengkap. Dengan adanya segmentasi konsumen yang baru, *P2P Lending* dapat mengembangkan rencana bisnis yang dimiliki dan dapat menarik minat banyak konsumen. Konsumen dan masyarakat yang tidak memiliki riwayat keuangan atau kredit juga dapat terbantu untuk mendapat akses pinjaman dengan adanya model penilaian kredit menggunakan fitur demografi sosial media.

## 5. KESIMPULAN

Berdasarkan hasil penelitian yang telah dilakukan, data *user demographic attributes* atau fitur demografi sosial media dapat dijadikan data penentu kelayakan kredit yang bisa menjadi data alternatif atau pengganti kekurangan data pada *P2P Lending*.

Dengan melakukan analisis dan pemahaman lebih, data sosial media yang tidak mampu menerjemahkan secara langsung konten terkait permasalahan kredit dapat digunakan. Pendekatan menggunakan teori prinsip 5C kredit juga perlu diperhatikan untuk mendapatkan kualitas data yang baik. Hasil dari *performance metrics* menunjukkan bahwa algoritma *random forest* memiliki kinerja tertinggi yang menghasilkan *accuracy* sebesar 81.25%, *precision* sebesar 87.88%, *recall* sebesar 63.04%, dan *f-measure* sebesar 73.42%, sedangkan algoritma *decision tree* menghasilkan *accuracy* sebesar 76.79%, *precision* sebesar 77.78%, *recall* sebesar 60.87%, dan *f-measure* sebesar 68.29%, sehingga algoritma *random forest* dipilih menjadi model penentu kelayakan kredit menggunakan fitur demografi sosial media yang terbaik. Dari hasil *performance metrics* yang ada, dapat diambil wawasan lain bahwa *P2P Lending* memiliki kesempatan untuk membuka segmentasi konsumen baru yang tidak memiliki riwayat keuangan atau kredit secara lengkap.

## 6. REFERENSI

- Amartha. (2019, Agustus 29). *Peer-to-peer Lending: Ini Keunggulan Fintech Lending*. Retrieved from Amartha Blog: <https://blog.amartha.com/ini-keunggulan-fintech-lending/>
- Breiman, L. (2001). Random Forests. *Machine Learning*, 5-32.
- DailySocial.id. (2019, Desember 4). *post: menggodok-data-membangun-credit-scoring*. Retrieved from DailySocial.id: <https://dailysocial.id/post/menggodok-data-membangun-credit-scoring>
- Fadillah, A. P. (2015). Penerapan Metode CRISP-DM untuk Prediksi Kelulusan Studi Mahasiswa Menempuh Mata Kuliah (Studi Kasus Universitas XYZ). *Jurnal Teknik Informatika dan Sistem Informasi* .
- Fahmi, I. (2014). *Manajemen Perkreditan*. Bandung: Alfabeta.
- Glassdoor. (2019). *Job: search*. Retrieved from Glassdoor.com: <https://www.glassdoor.com/Job/Home/recentActivity.htm>
- Guo, G., Zhu, F., Chen, E., Liu, Q., Wu, L., & and Guan, C. (2016). From footprint to evidence: An exploratory study of mining social data for credit scoring. *ACM Transactions on the Web*, 10(4), 22:1-38. doi:10.1145/2996465

- Hua, W., Wang, Z., Wang, H., Zheng, K., & Zhou, X. (2015). Short Text Understanding Through Lexical-Semantic Analysis. *International Conference on Data Engineering (ICDE)*, (pp. 495-506).
- Ikatan Bankir Indonesia. (2015). *Bisnis Kredit Perbankan*. Jakarta Pusat: Gramedia Pustaka Utama.
- Kelly and Persol Indonesia. (2018). *2018 Salary Guide*.
- Koinworks.com (2016, Oktober). *Blog: Apa itu peer to peer Lending (P2P Lending)? Cari tahu selengkapnya!*. Retrieved from Koinworks.com: [koinworks.com/blog/ketahui-tentang-peer-peer-lending/](http://koinworks.com/blog/ketahui-tentang-peer-peer-lending/)
- Michael Page. (2020). *Indonesian Salary Benchmark 2020*.
- Otoritas Jasa Keuangan. (2020). *Perkembangan Fintech Lending Periode Januari 2020*.
- Refaeilzadeh, P., Tang, L., & Liu, H. (2009). Cross-Validation. In: LIU L., ÖZSU M.T. (eds) Encyclopedia of Database Systems. Boston: Springer.
- Suyanto. (2019). *Data Mining Untuk Klasifikasi dan Klasterisasi Data, Edisi Revisi*. Bandung: Informatika.
- Wibowo, A.T., Saikhu, A., & Soelaiman, R. (2016). Implementasi Algoritma Deteksi Spam yang Tersisipi Informasi Citra dengan Metode SVM dan Random Forest (Tugas Akhir, Institut Teknologi Sepuluh November).