

# *Digital Awareness* Perbankan: Analisis Kinerja Perbankan Pascapandemi COVID-19 Berbasis Teks Laporan Tahunan

Okiriza Wibisono\* dan Ali Akbar Septiandri

\*E-mail: [okiriza.w@gmail.com](mailto:okiriza.w@gmail.com)

## Abstrak

Pandemi COVID-19 telah berdampak pada berbagai sektor sosial-ekonomi di seluruh dunia, termasuk sektor perbankan. Bank dituntut untuk dapat menyediakan layanan digital dan berkompetisi di ranah digital termasuk dengan fintech, di tengah pembatasan mobilitas yang berdampak pada penurunan aktivitas ekonomi dan kinerja keuangan debitur. Digitalisasi perbankan menjadi salah satu solusi untuk menjawab tantangan tersebut. Penelitian ini memaparkan metodologi alternatif untuk mengukur tingkat *digital awareness* sebagai indikator digitalisasi bank. Memanfaatkan metode analisis teks / *natural language processing* (NLP), indikator-indikator *digital awareness* bank dapat dibentuk secara otomatis dari jumlah kemunculan kata-kata terkait digitalisasi dalam laporan tahunan bank. Hasil pengujian model regresi pada data 47 emiten bank di Indonesia menunjukkan indikator tersebut tidak signifikan memprediksi kinerja bank dalam jangka pendek pascapandemi COVID-19. Namun, mempertimbangkan kemudahannya untuk diterapkan, metodologi yang diajukan berpotensi untuk dikembangkan lebih lanjut oleh otoritas atau peneliti, antara lain untuk melengkapi *hard indicators* digitalisasi perbankan, untuk analisis dampak digitalisasi dalam jangka panjang dan/atau terhadap variabel perbankan lainnya, ataupun untuk analisis dampak digitalisasi pada sektor keuangan secara lebih luas.

Kata kunci: digitalisasi perbankan, pandemi COVID-19, kinerja bank, analisis teks

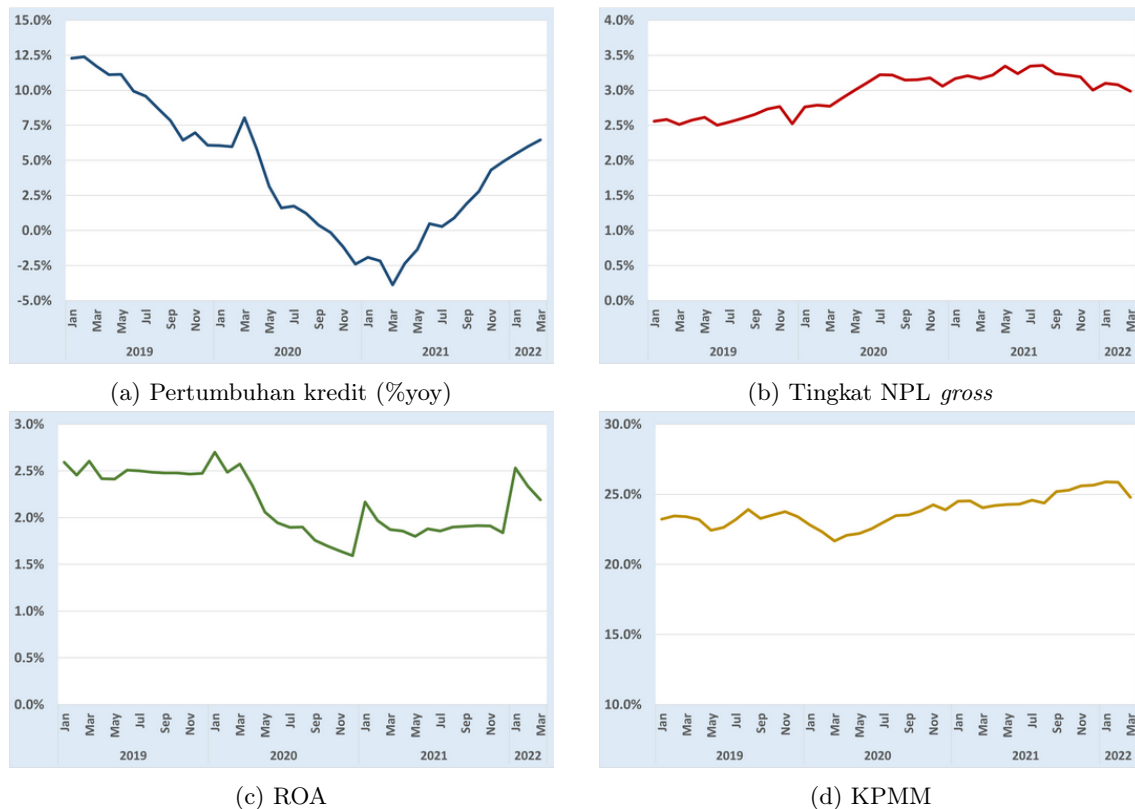
JEL code: Banks, Depository Institutions, Micro Finance Institutions, Mortgages (G21); Firm Performance: Size, Diversification, and Scope (L25); Technological Change: Choices and Consequences, Diffusion Processes (O33)

## 1 Latar Belakang dan Tujuan

Pandemi COVID-19 telah berdampak pada tatanan sosial-ekonomi di seluruh dunia, termasuk pada sektor perbankan, sebagaimana ditunjukkan dalam berbagai penelitian. Pembatasan mobilitas mengakibatkan penurunan aktivitas ekonomi yang berdampak kepada kinerja korporasi. Perbankan menjadi salah satu sektor yang memegang peranan penting dalam mengatasi dampak pandemi COVID-19, antara lain melalui implementasi kebijakan restrukturisasi kredit yang dikeluarkan oleh Otoritas Jasa Keuangan (OJK). Industri perbankan Indonesia secara agregat memiliki tingkat ketahanan yang cukup baik, di antaranya angka kewajiban penyediaan modal minimum (KPMM) yang stabil dan bahkan meningkat di era pandemi: dari 23,4% di akhir 2019, meningkat menjadi 23,89% di 2020 dan 25,67% di 2021.

Di sisi lain, pandemi COVID-19 juga berdampak terhadap risiko dan kinerja perbankan. Terjadi peningkatan risiko kredit perbankan yang tecermin dari meningkatnya rasio *non-performing loan* (NPL) dari sekitar 2,53% di akhir 2019 hingga mencapai 3,35% di Agustus 2021, dan sampai dengan pertengahan 2022 masih berada di kisaran 3% (sumber: Statistik Perbankan Indonesia - OJK). Selain itu, kinerja/rentabilitas perbankan yang diukur dengan *return on assets* (ROA) secara agregat turun cukup signifikan dari 2,47% di 2019 menjadi 1,59% di 2020 dan sedikit mengalami pemulihan ke 1,84% pada akhir 2021. Ilustrasi lebih lengkap dapat dilihat pada Gambar 1.

Di samping itu, pembatasan mobilitas dan kontak fisik juga menuntut perbankan untuk menyediakan layanan digital bagi nasabahnya. Bank yang tidak mampu berkompetisi dan/atau berintegrasi dengan ekosistem digital berisiko kehilangan market share-nya, baik kepada bank lain yang lebih siap secara digital maupun kepada pemain baru seperti *fintech* ([Boston Consulting Group, 2020](#)).



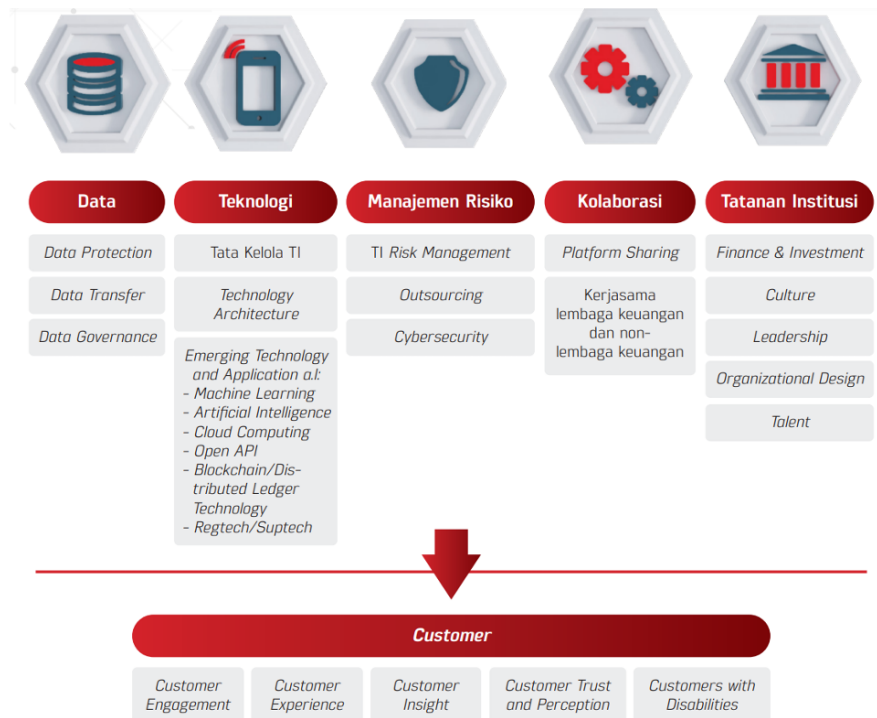
Gambar 1: Perkembangan indikator perbankan agregat

Penelitian ini mengkaji pengaruh digitalisasi perbankan terhadap kinerja perbankan. Secara spesifik, dilakukan analisis regresi yang membandingkan kinerja bank yang telah/sedang melakukan transformasi digital sebelum pandemi COVID-19, dengan bank yang relatif belum bertransformasi digital. Tingkat “*digital awareness*” tersebut diukur menggunakan metode analisis teks terotomasi dari pemaparan bank dalam laporan tahunannya. Indikator *digital awareness* yang diperoleh dari analisis teks tersebut memiliki keunggulan yaitu dapat dibentuk dengan mudah dan cepat, relatif jika dibandingkan dengan pengumpulan data terkait digitalisasi perbankan antara lain melalui survei dan pelaporan. Hasil pengujian dengan model regresi menunjukkan indikator tersebut tidak signifikan dalam memprediksi kinerja bank dalam jangka pendek (periode pra- dan pascapandemi COVID-19). Namun mengingat digitalisasi cenderung memberikan dampak dalam jangka panjang, dapat dilakukan penelitian lebih lanjut untuk menguji indikator *digital awareness* tersebut.

Kontribusi yang diberikan oleh penelitian ini yaitu:

1. memaparkan metodologi alternatif untuk mengukur tingkat digital awareness perbankan berdasarkan publikasinya, khususnya laporan tahunan;
2. menganalisis tingkat digital awareness emiten perbankan di Indonesia sebelum pandemi COVID-19; dan
3. menganalisis pengaruh digital awareness perbankan sebelum pandemi COVID-19 terhadap kinerjanya pascapandemi COVID-19.

Struktur karya tulis ini selanjutnya sebagai berikut. Bab 2 membahas studi literatur yang terbagi ke dalam studi literatur terkait digitalisasi perbankan dan studi literatur terkait analisis teks/ *natural language processing* (NLP). Bab 3 memaparkan data dan metodologi yang digunakan untuk membentuk indikator *digital awareness* bank. Bab 4 menyajikan hasil dan analisis pengujian indikator tersebut. Kesimpulan penelitian dan rekomendasi kebijakan disampaikan dalam Bab 5.



Gambar 2: Cetak Biru Transformasi Digital Perbankan (Otoritas Jasa Keuangan, 2021)

## 2 Studi Literatur

### 2.1 Pandemi COVID-19, Digitalisasi, dan Kinerja Perbankan

Berbagai penelitian telah mempelajari dampak pandemi COVID-19 terhadap sektor perbankan. Berger et al. (2021) merangkum isu-isu yang dihadapi perbankan dan otoritas dalam era pandemi COVID-19 dan periode pascapandemi. Isu yang mengemuka antara lain ketahanan perbankan, intermediasi perbankan, risiko kredit, dan evaluasi dampak kebijakan otoritas. Salah satu implikasi atau area penelitian lebih lanjut yang diidentifikasi oleh Berger et al. adalah dampak pandemi COVID-19 terhadap digitalisasi dan model bisnis bank. Penting untuk diketahui apakah digitalisasi dapat meningkatkan ketahanan dan kinerja bank dalam periode tekanan.

Penelitian mengenai topik serupa telah dilakukan oleh Assaf et al. (2019). Menggunakan data perbankan di AS pada periode 1986-2009 (sebelum pandemi COVID-19), mereka menemukan bahwa efisiensi biaya di periode normal dapat membantu ketahanan dan kinerja bank dalam periode krisis. Penelitian mengenai perbankan Indonesia dilakukan antara lain oleh Syamsudin (2022). Indikator-indikator digitalisasi berupa volume transaksi di berbagai channel perbankan seperti ATM, mobile banking, dan internet banking, dapat membantu prediksi return on assets untuk studi kasus bank-bank besar. Adapun studi mengenai kinerja (profitabilitas) bank secara makro di emerging markets telah dipelajari oleh Dietrich and Wanzenried (2014) dan Kohlscheen et al. (2018). Beberapa variabel makro yang mempengaruhi kinerja bank yaitu pertumbuhan kredit, suku bunga jangka panjang, dan *sovereign risk*. Sedangkan faktor signifikan yang bersifat bank-specific antara lain rasio biaya terhadap pendapatan (*cost-to-income*), cadangan kerugian penurunan nilai (CKPN), share pendapatan bunga, tingkat biaya bunga, dan kepemilikan bank (negara atau asing).

Jika dikaitkan dengan premis bahwa digitalisasi dapat meningkatkan efisiensi bank (Otoritas Jasa Keuangan, 2021), maka dapat diuji hipotesis bahwa bank yang sudah lebih jauh bertransformasi digital, secara rata-rata akan memiliki tingkat ketahanan dan/atau kinerja yang lebih baik. Cetak Biru Transformasi Digital Perbankan yang disusun oleh OJK telah mengidentifikasi 5 (lima) faktor utama dalam digitalisasi perbankan: data, teknologi, manajemen risiko, kolaborasi, dan tatanan institusi, sebagaimana pada Gambar 2.

Yusgiantoro (2018) juga secara spesifik telah meneliti dampak adopsi teknologi digital terhadap kinerja perbankan Indonesia. Digitalisasi didefinisikan sebagai share biaya yang terkait IT terhadap

total biaya operasional. Indikator digitalisasi tersebut diuji terhadap 3 (tiga) indikator bank, yaitu intermediasi, kinerja, dan jangkauan pasar (*market outreach*). Kesimpulan yang dihasilkan cukup menarik: digitalisasi dapat meningkatkan jangkauan pasar, namun digitalisasi yang terlalu agresif berdampak negatif pada kinerja. Sejalan dengan hasil penelitian tersebut, [Beccalli \(2007\)](#) juga menunjukkan hubungan yang lemah antara investasi IT dengan kinerja dan efisiensi bank, yang disebutnya suatu “profitability paradox”.

Sejalan dengan penelitian-penelitian di atas dan Cetak Biru Transformasi Digital Perbankan tersebut, penelitian ini menguji apakah indikator digitalisasi bank berdampak pada kinerja perbankan, khususnya digitalisasi sebelum pandemi COVID-19 (2018-2019) dan dampaknya pada kinerja pascapandemi (2021). Indikator digitalisasi yang digunakan adalah penulisan kata kunci terkait sejumlah faktor pada Gambar 2 dalam laporan tahunan bank.

## 2.2 Analisis Teks (NLP)

Dalam satu dekade terakhir, pemanfaatan *big data analytics* berkembang pesat karena didorong oleh peningkatan jumlah data secara signifikan (antara lain yang bersumber dari *smartphone*, *e-commerce*, dan *fintech*), peningkatan teknologi untuk mengumpulkan dan mengolah data (antara lain *big data tools*, *cloud computing*), serta penemuan dan aplikasi berbagai algoritme *machine learning* ([Financial Stability Board, 2017](#)).

Salah satu pemanfaatan *big data analytics* di sektor keuangan adalah *natural language processing* (NLP). NLP atau *text mining* adalah sekumpulan metode untuk mengolah/menganalisis data teks yang bersifat tidak terstruktur hingga membentuk variabel-variabel atau informasi yang bersifat terstruktur dan dapat diproses lebih lanjut. Sebagai contoh, dari kalimat “Perubahan perilaku nasabah dan adopsi masyarakat yang tinggi terhadap perkembangan teknologi digital mendorong perbankan terus beradaptasi serta melakukan inovasi tiada henti” pada laporan tahunan salah satu bank, dapat dilakukan analisis sentimen untuk mengekstraksi tonasi positif dari bank tersebut, khususnya terkait inovasi/teknologi.

Dalam literatur telah terdapat berbagai penelitian yang memanfaatkan NLP di sektor keuangan. Beberapa contohnya sebagai berikut. [Bollen et al. \(2011\)](#) membuktikan bahwa sentimen yang diekspresikan pengguna Twitter dapat dianalisis untuk meningkatkan akurasi prediksi harga saham. Sentimen atau *mood* tersebut diidentifikasi dengan kata kunci seperti “I feel ...” dan “I’m ...”, dan kemudian dikelompokkan ke dalam kategori-kategori berdasarkan OpinionFinder dan Google-Profile of Mood States (GPOMS). Serupa dengan [Bollen et al.](#), [Carlini et al. \(2020\)](#) melakukan analisis terhadap konten pemberitaan perbankan dan menemukan bahwa pemberitaan negatif berdampak signifikan terhadap harga saham bank. Salah satu penelitian yang paling dikenal adalah indikator Economic Policy Uncertainty (EPU) oleh [Baker et al. \(2016\)](#). Dengan menggunakan kemunculan sejumlah kata kunci terkait ekonomi, kebijakan, dan ketidakpastian di pemberitaan, dibentuklah indikator EPU yang telah menjadi salah satu indikator makro yang diacu oleh berbagai pengambil kebijakan di seluruh dunia.

Literatur yang paling mendekati penelitian ini antara lain dipublikasikan oleh [Loughran and McDonald \(2011\)](#). [Loughran and McDonald](#) menguji dampak penulisan kata kunci negatif pada laporan tahunan emiten di AS (*10-K filing*) terhadap berbagai indikator seperti kinerja/*return*, volume perdagangan saham, volatilitas harga saham, *fraud*, dan penghasilan di luar ekspektasi. [Loughran and McDonald \(2016\)](#) merangkum berbagai penerapan NLP di sektor keuangan serta teknik yang juga digunakan dalam penelitian ini, antara lain *word lists* (daftar kata kunci), matriks frekuensi kata kunci, metode pengelompokan kata kunci/variabel seperti algoritme *latent semantic analysis* (LSA), dan analisis sentimen.

## 3 Data dan Metodologi

### 3.1 Data

Penelitian ini menggunakan data emiten perbankan yang tercatat di Bursa Efek Indonesia (BEI), dengan periode pengamatan tahun 2018-2019 (sebelum pandemi COVID-19) dan 2021 (pascapandemi COVID-19). Terdapat 47 emiten dalam periode pengamatan.

Jenis data yang digunakan yaitu:

1. **Laporan keuangan bank:** untuk indikator *return on assets* (ROA), aset, *loan-to-deposit ratio* (LDR), *net interest margin* (NIM), dan rasio biaya operasional terhadap pendapatan operasional (BOPO). Indikator-indikator tersebut berperan sebagai variabel dependen dan variabel kontrol dalam model regresi.
2. **Laporan tahunan bank:** untuk indikator jumlah kemunculan kata kunci terkait digitalisasi bank yang dimuat dalam laporan tahunan 2018. Indikator-indikator tersebut berperan sebagai variabel yang diuji (*variables of interest*) dalam model regresi.

Kedua jenis data tersebut diperoleh dari situs BEI dan/atau situs masing-masing bank. Adapun variabel makro (a.l. pertumbuhan ekonomi, suku bunga, dan kinerja industri perbankan secara agregat) tidak digunakan mengingat analisis regresi yang dilakukan bersifat *cross section* antara 2 periode waktu, yakni membandingkan dampak *digital awareness* bank sebelum pandemi COVID-19 terhadap kinerjanya pascapandemi COVID-19.

Untuk laporan tahunan (sumber indikator digital awareness), periode yang digunakan adalah laporan tahunan 2018, bukan laporan tahunan 2019. Hal ini untuk mengurangi bias terkait pandemi COVID-19 dalam penulisan laporan tahunan bank, mengingat batas waktu penyampaian laporan tahunan 2019 oleh emiten perbankan adalah pada Juni 2020 (sudah memasuki beberapa bulan sejak pandemi COVID-19 di Indonesia).

## 3.2 Metodologi

### 3.2.1 Ekstraksi Data Teks

Sebagaimana dijelaskan secara singkat dalam subbagian 3.1, data laporan tahunan bank diolah untuk membentuk sejumlah indikator *digital awareness* bank. Indikator tersebut dibentuk dengan menghitung jumlah kemunculan kata kunci tertentu dalam laporan tahunan bank. Daftar kata kunci lengkap dan variasi penulisannya dapat dilihat pada Lampiran A.

Tahapan pengolahan yang dilakukan pada data laporan tahunan yaitu:

1. Membuang bab laporan keuangan secara manual (bagian yang berupa tabel-tabel neraca, laba rugi, laporan arus kas, laporan ekuitas, dan lain-lain).
2. Ekstraksi teks dari file PDF laporan tahunan<sup>1</sup>.
3. Menghitung jumlah kemunculan masing-masing kata kunci dalam teks laporan tahunan<sup>2</sup>. Keluaran dari tahap ini adalah matriks frekuensi kata kunci (TF), dengan dimensi bank  $\times$  kata kunci (jumlah kemunculan kata kunci  $Y$  pada laporan tahunan 2018 bank  $X$ ).
4. Membentuk sejumlah indikator digital awareness dari matriks frekuensi kata kunci. Keluaran dari tahap ini adalah matriks digital awareness, dengan dimensi bank  $\times$  indikator, di mana jumlah indikator  $\leq$  jumlah kata kunci (dilakukan suatu pengelompokan indikator). Matriks ini menjadi **variabel yang diuji** dalam regresi.

Ilustrasi dari proses ini dapat dilihat pada Gambar 3.

Adapun jumlah kemunculan kata kunci tidak dapat digunakan secara langsung sebagai variabel regresi mengingat terdapat  $> 100$  kata kunci, sementara jumlah bank yang dianalisis hanya 47 ( $n < p$ ). Oleh karena itu dilakukan pengelompokan sebagaimana poin 4.

Terdapat sejumlah teknik yang dapat digunakan untuk mengelompokkan/membentuk indikator *digital awareness* dari matriks frekuensi kata kunci. Teknik yang digunakan dalam penelitian ini adalah *latent semantic analysis* (LSA) (Landauer et al., 1998) dan k-means clustering (Bishop and Nasrabadi, 2006, hal. 424–428). Selain pengelompokan berdasarkan algoritme, juga dilakukan pengelompokan kata kunci secara manual sebagaimana Lampiran A, tetapi hasil/kesimpulan uji signifikansi yang didapat tidak berubah.

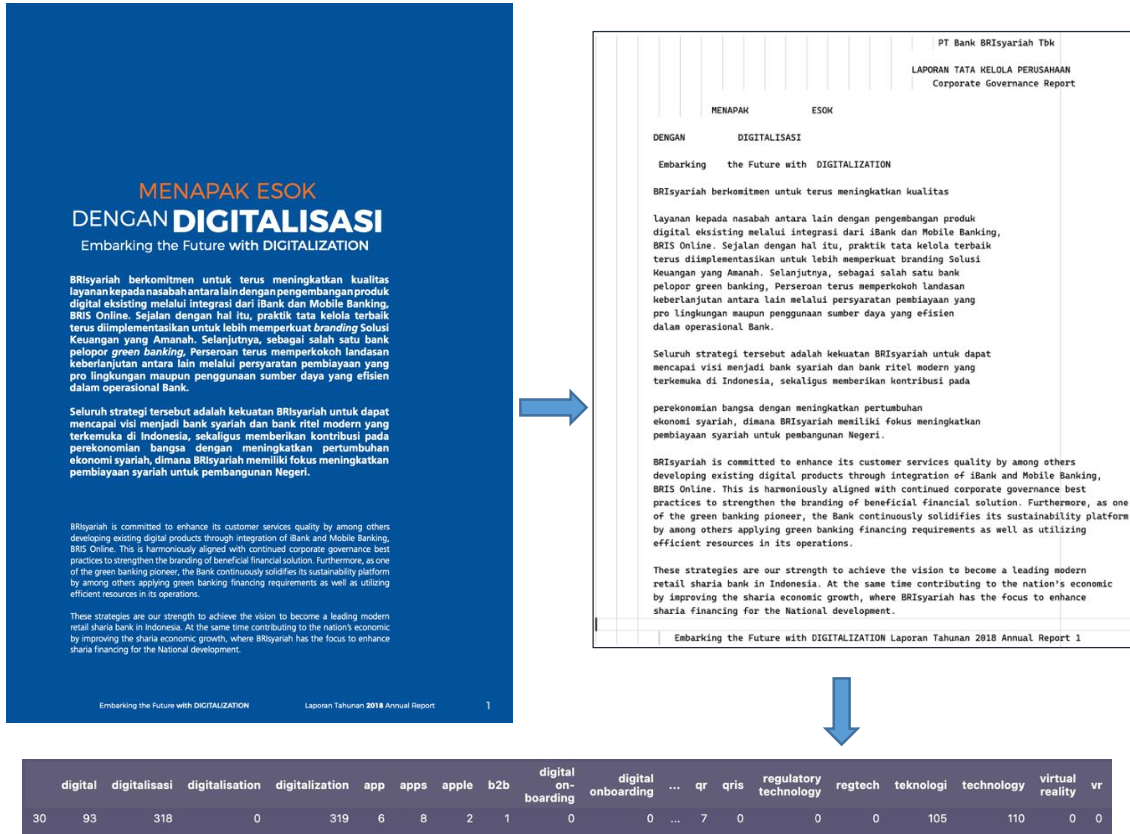
### 3.2.2 Regresi

Berikut model regresi yang digunakan untuk menguji hipotesis dampak digital awareness terhadap kinerja perbankan pascapandemi COVID-19.

<sup>1</sup>Menggunakan pustaka Python pdfplumber (<https://github.com/jsvine/pdfplumber>)

<sup>2</sup>Menggunakan modul CountVectorizer dalam pustaka Python scikit-learn (<https://scikit-learn.org/>)





Gambar 3: Proses ekstraksi dari PDF laporan tahunan bank menjadi teks, lalu menjadi matriks frekuensi kata kunci (tahapan 1-3). Seluruh proses dilakukan secara otomatis.

$$\Delta \widehat{ROA}_{i,t,t-1} = \beta_0 + \sum_k \beta_k \text{kontrol}_{i,k,t-1} + \sum_j \beta_j \text{digital awareness}_{i,j,t-1} \quad (1)$$

Penjelasan persamaan 1 di atas adalah sebagai berikut:

1. **Variabel dependen** yang diestimasi adalah perubahan *return on assets* (ROA) (selisih sesudah pandemi dan sebelum pandemi). Penggunaan delta ROA tersebut bertujuan menormalisasi nilai ROA, mengingat terdapat sejumlah bank yang memiliki level ROA lebih tinggi dibandingkan bank lainnya. Selain itu, penggunaan delta ROA juga meminimalisasi kemungkinan *false positive* mengingat delta ROA lebih sulit diprediksi dibandingkan level ROA.
2. **Periode data** yang digunakan adalah:  
 $t$  = tahun 2021 (khusus variabel dependen / ROA)  
 $t-1$  = rata-rata 2018 dan 2019 untuk variabel dependen dan kontrol (dari laporan keuangan), dan 2018 untuk variabel digital awareness (dari teks laporan tahunan)
3.  $i$  = indeks untuk bank (terdapat 47 bank/emiten dalam analisis)
4. **Variabel kontrol** terdiri dari (log) aset, *loan-to-deposit ratio* (LDR), *net interest margin* (NIM), dan rasio biaya operasional terhadap pendapatan operasional (BOPO). Nilai yang diambil adalah rata-rata tahun 2018 dan 2019.
5. **Variabel yang diuji** (*variables of interest*) terdiri dari sekumpulan indikator *digital awareness* yang diekstraksi dari teks laporan tahunan bank tahun 2018, menggunakan metodologi sebagaimana dijelaskan dalam subbagian 3.2.1.

Persamaan regresi di atas diestimasi dengan metode *ordinary least squares* (OLS) mengingat data yang digunakan bersifat *cross section*. Metode *robust regression* juga telah dicoba, tetapi hasil/kesimpulan yang diperoleh tidak banyak berubah.

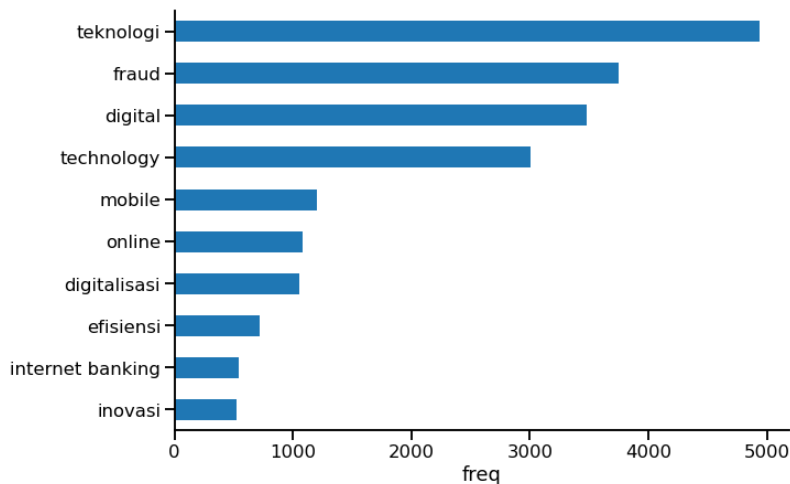
Tabel 1: Deskripsi variabel dependen dan variabel kontrol

Variabel		Penjelasan	Perhitungan
Dependen (Y)	<b>delta_roa</b>	Kinerja bank yang diukur dari <i>return on assets</i> (ROA)	ROA tahun 2021 – rata-rata ROA tahun 2018-2019
Kontrol	<b>asset</b>	Ukuran bank berupa jumlah aset	log(rata-rata aset tahun 2018-2019)
	<b>ldr</b>	Rasio kredit yang diberikan terhadap dana pihak ketiga ( <i>loan-to-deposit ratio</i> )	rata-rata LDR tahun 2018-2019
	<b>nim</b>	<i>Net interest margin</i> (NIM)	rata-rata NIM tahun 2018-2019
	<b>bopo</b>	Rasio biaya operasional terhadap pendapatan operasional	rata-rata BOPO tahun 2018-2019

## 4 Hasil Analisis

### 4.1 Variabel Teks

Dengan metodologi ekstraksi data teks sebagaimana dipaparkan dalam subbagian 3.2.1, diperoleh 10 kata kunci yang paling banyak disebutkan dalam laporan tahunan bank tahun 2018 seperti pada Gambar 4.

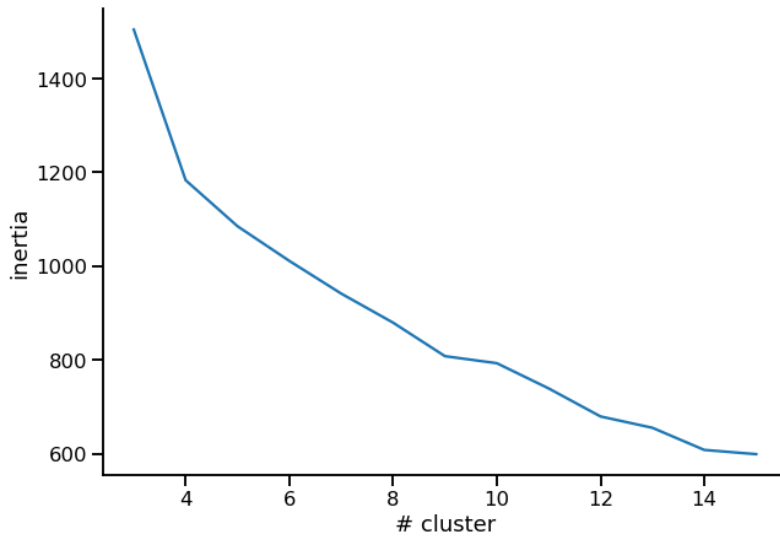


Gambar 4: Sepuluh kata kunci yang paling sering muncul dari semua laporan tahunan bank

Salah satu pengamatan dari hasil di atas adalah terdapat kata kunci yang ditulis dalam bahasa Indonesia dan bahasa Inggris, antara lain “teknologi” dan “technology”. Untuk itu, diperlukan suatu pengelompokan kata kunci dalam membentuk indikator *digital awareness*, baik pengelompokan secara manual maupun menggunakan teknik tertentu a.l. *k-means clustering* dan LSA.

Dengan *k-means clustering*, diperoleh sejumlah kelompok kata kunci berdasarkan *co-occurrence*-nya, yaitu pola kemunculannya di berbagai laporan tahunan bank. Gambar 5 menunjukkan grafik *scree plot* jumlah kelompok kata kunci berdasarkan metrik *inertia* (Bishop and Nasrabadi, 2006, hal. 427–428). Berdasarkan grafik tersebut, dapat dilihat bahwa penurunan nilai *inertia* terbesar terjadi saat jumlah klaster berubah dari tiga menjadi empat. Terlihat pula bahwa perubahan nilai *inertia* dari empat menjadi lima klaster menjadi lebih kecil. Metode pencarian jumlah klaster terbaik ini dikenal juga sebagai *elbow method* atau metode siku. Analisis lebih lanjut menunjukkan bahwa semakin banyak klaster yang digunakan, maka anggota klaster akan semakin sedikit. Sebagai contoh, meski nilai *inertia* saat jumlah klaster 15 adalah yang paling kecil, tetapi terdapat 7 klaster yang hanya memiliki satu anggota.

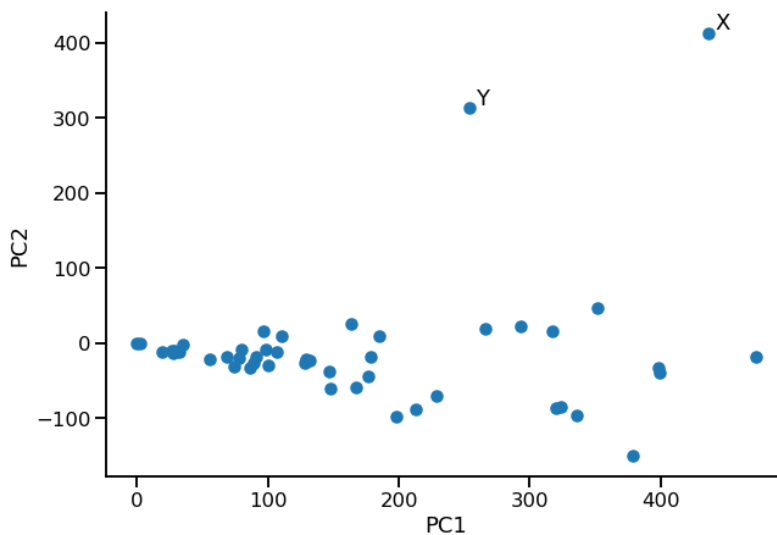
Teknik lain yang dapat digunakan untuk mengelompokkan kata kunci adalah *latent semantic analysis* (LSA). LSA dilakukan dengan menerapkan *singular value decomposition* (SVD) pada matriks frekuensi kata kunci. Ditemukan bahwa 6 (enam) *principal components* dapat menjelaskan 97,6% variansi dalam matriks frekuensi kata kunci tersebut. Keenam *principal components* ini dapat



Gambar 5: Nilai *inertia* dari *k-means* dengan berbagai jumlah kluster

digunakan sebagai variabel dalam regresi.

Dengan hanya menggunakan dua *principal components*, dapat dilihat bahwa terdapat dua pencilan dalam data berdasarkan *principal component* kedua seperti yang divisualisasikan dalam Gambar 6. Analisis lebih lanjut menunjukkan bahwa dua bank yang menjadi pencilan tersebut sering menggunakan kata kunci “digitalisasi” dalam laporan tahunannya. Terdapat dua bank yang menggunakan kata kunci “digitalisasi” 318 dan 479 kali, secara berurutan, dalam laporan tahunan di tahun 2018. Jumlah ini 10 sampai 15 kali lebih banyak dibandingkan bank ketiga yang paling sering menggunakan kata “digitalisasi” dengan hanya 32 kali kemunculan.



Gambar 6: Proyeksi data pada dua *principal components* pertama dari matriks frekuensi kata kunci

## 4.2 Regresi

Hasil uji signifikansi model dasar regresi (tanpa indikator *digital awareness*) menunjukkan bahwa variabel kontrol yang digunakan tidak signifikan dalam memprediksi kinerja bank, sebagaimana Tabel 2.

Selanjutnya, untuk menguji signifikansi indikator *digital awareness*, model dasar regresi tersebut ditambahkan 6 (enam) *principal components* sebagai prediktor tambahan untuk menjelaskan ki-



Tabel 2: Hasil uji signifikansi model dasar (hanya variabel kontrol)  
 $R^2 = 0.033$ , adjusted  $R^2 = -0.059$

	coef	std err	t	P> t	[0.025	0.975]
intercept	-0.0698	0.073	-0.961	0.342	-0.216	0.077
ldr	0.0049	0.012	0.399	0.692	-0.020	0.030
nim	-0.1019	0.221	-0.462	0.647	-0.548	0.344
bopo	0.0347	0.034	1.020	0.314	-0.034	0.103
log(assets)	0.0025	0.004	0.608	0.546	-0.006	0.011

nerja bank, sebagaimana dipaparkan dalam subbagian 4.1. Hasil pengujian dapat terlihat pada Tabel 3.

Tabel 3: Hasil uji signifikansi indikator *digital awareness*  
 $R^2 = 0.105$ , adjusted  $R^2 = -0.143$

	coef	std err	t	P> t	[0.025	0.975]
intercept	-0.0433	0.099	-0.438	0.664	-0.244	0.157
ldr	0.0055	0.013	0.421	0.676	-0.021	0.032
nim	-0.2094	0.266	-0.787	0.436	-0.749	0.330
bopo	0.0388	0.039	1.001	0.324	-0.040	0.118
log(assets)	0.0035	0.008	0.425	0.673	-0.013	0.020
pc1	-0.0032	0.003	-0.987	0.330	-0.010	0.003
pc2	0.0038	0.004	0.898	0.375	-0.005	0.012
pc3	0.0045	0.005	0.941	0.353	-0.005	0.014
pc4	0.0043	0.005	0.876	0.387	-0.006	0.014
pc5	0.0019	0.005	0.366	0.717	-0.009	0.013
pc6	-0.0016	0.006	-0.254	0.801	-0.014	0.011

Beberapa pengamatan yang dapat disimpulkan dari Tabel 2 dan Tabel 3 antara lain:

1. Koefisien intercept bernilai negatif, yang mencerminkan penurunan ROA perbankan dari periode sebelum pandemi COVID-19 (2018-2019) ke periode pascapandemi (2021).
2. Indikator *digital awareness* (keenam *principal components*) tidak signifikan secara statistik dalam menjelaskan perubahan ROA. Hal ini tecermin dari nilai *p-value* paling kecil 0.33 dan *adjusted R<sup>2</sup>* berkurang dari  $-0.059$  pada model dasar menjadi  $-0.143$ .
3. Arah/tanda koefisien pada variabel kontrol juga di luar ekspektasi (contoh: NIM memiliki koefisien negatif terhadap ROA).
4. Telah diuji pula metode *robust regression* mengingat terdapat pencilan (*outliers*) dalam data ROA yang diobservasi. Namun, hasil *robust regression* tersebut (Lampiran B) tidak mengubah kesimpulan.

## 5 Kesimpulan dan Rekomendasi Kebijakan

### 5.1 Kesimpulan

Penelitian ini bertujuan menguji dampak digitalisasi bank sebelum pandemi COVID-19 (2018-2019) terhadap kinerja bank setelah pandemi (2021). Hal ini didasarkan pada hipotesis umum dan sejumlah penelitian yang memprediksi dampak positif digitalisasi terhadap kinerja bank, antara lain melalui efisiensi biaya.

Dalam penelitian ini tingkat digitalisasi bank diukur berdasarkan kemunculan kata-kata terkait digitalisasi dalam laporan tahunan bank tahun 2018 (sebelum pandemi COVID-19). Menggunakan metode analisis teks / *natural language processing* (NLP), secara otomatis dapat dibentuk sejumlah indikator *digital awareness* dari jumlah kemunculan kata tersebut, yang sebagian didasarkan

pada faktor-faktor digitalisasi dalam Cetak Biru Digitalisasi Perbankan (Otoritas Jasa Keuangan, 2021). Indikator *digital awareness* yang dihasilkan kemudian diuji dalam model regresi kinerja bank sebelum dan setelah pandemi (selisih *return on assets* 2018-2019 dan 2021).

Hasil penelitian ini menunjukkan bahwa indikator *digital awareness* yang diekstraksi dari laporan tahunan bank tersebut tidak dapat membantu memprediksi perubahan kinerja bank antara sebelum dan pascapandemi COVID-19 (tidak signifikan secara statistik). Dengan kata lain, bank-bank yang *digital aware* dalam artian banyak menyebutkan kata kunci terkait digitalisasi dalam laporan tahunannya, memiliki tren kinerja pascapandemi yang secara umum tidak berbeda dengan bank lainnya. Hasil tersebut cukup sejalan dengan penelitian Yusciantoro (2018) dan Beccalli (2007) yang mengatakan bahwa biaya/investasi IT tidak berdampak positif pada kinerja keuangan bank, meskipun terbukti dapat meningkatkanjangkauan pasar.

## 5.2 Rekomendasi Kebijakan

Dari kesimpulan di atas dapat disampaikan sejumlah rekomendasi untuk penelitian lebih lanjut dan untuk kebijakan sektor perbankan:

1. Otoritas/peneliti dapat memanfaatkan metodologi yang diajukan untuk membentuk indikator *digital awareness* yang **melengkapi *hard indicators*** digitalisasi suatu bank (*hard indicators* antara lain jumlah pengguna *e-banking*, volume transaksi *e-banking*, dan implementasi layanan/teknologi/channel tertentu). Otoritas/peneliti juga dapat menggunakan daftar kata kunci dan pengelompokan lainnya yang sesuai kebutuhan.
2. Penelitian ini menggunakan perubahan ROA sebagai variabel dependen. Dapat dilakukan penelitian lebih lanjut menggunakan **variabel dependen lainnya** antara lain ketahanan, intermediasi, market share, return harga saham, atau kombinasi-kombinasinya. Selain itu, terdapat potensi digitalisasi memberikan dampak positif kepada bank dalam **jangka panjang** misalnya menggunakan **metode panel**, sehingga dapat dilakukan penelitian lebih lanjut dengan periode data yang lebih panjang khususnya di era pascapandemi COVID-19 ini.
3. Hasil pengujian indikator *digital awareness* yang tidak signifikan terhadap kinerja bank dapat dipandang sebagai suatu hal yang positif bagi bank yang memiliki tingkat *digital awareness* di bawah rata-rata. Hal ini mengingat tingkat *digital awareness* tersebut tidak berkorelasi dengan kinerja (dalam hal ini ROA) pascapandemi COVID-19. Dengan demikian, bank-bank tersebut masih berkesempatan untuk melakukan digitalisasi antara lain untuk mempertahankan/meningkatkan *market share* pada segmen tertentu.
4. Bagi otoritas yang memiliki data yang lebih lengkap, metodologi yang diajukan dapat diuji pada sektor **perbankan secara lebih luas**, tidak hanya bank yang merupakan emiten bursa. Untuk itu, otoritas dapat memanfaatkan **dokumen lain** dari bank seperti dokumen rencana bisnis bank. Otoritas juga dapat mencoba menerapkan metodologi ini pada **sektor lainnya** seperti BPR maupun sektor keuangan secara umum.

## Daftar Referensi

- Assaf, A. G., Berger, A. N., Roman, R. A., and Tsionas, M. G. (2019). Does efficiency help banks survive and thrive during financial crises? *Journal of Banking & Finance*, 106:445–470.
- Baker, S. R., Bloom, N., and Davis, S. J. (2016). Measuring economic policy uncertainty. *The quarterly journal of economics*, 131(4):1593–1636.
- Beccalli, E. (2007). Does it investment improve bank performance? evidence from europe. *Journal of Banking & Finance*, 31(7):2205–2230.
- Berger, A. N., Demirgüç-Kunt, A., Moshirian, F., and Saunders, A. (2021). The way forward for banks during the covid-19 crisis and beyond: Government and central bank responses, threats to the global banking industry. *Journal of Banking & Finance*, 133:106303.
- Bishop, C. M. and Nasrabadi, N. M. (2006). *Pattern Recognition and Machine Learning*, volume 4. Springer.

- Bollen, J., Mao, H., and Zeng, X. (2011). Twitter mood predicts the stock market. *Journal of computational science*, 2(1):1–8.
- Boston Consulting Group (2020). The rise of digital banking in southeast asia.
- Carlini, F., Cucinelli, D., Previtali, D., and Soana, M. G. (2020). Don’t talk too bad! stock market reactions to bank corporate governance news. *Journal of Banking & Finance*, 121:105962.
- Dietrich, A. and Wanzenried, G. (2014). The determinants of commercial banking profitability in low-, middle-, and high-income countries. *The Quarterly Review of Economics and Finance*, 54(3):337–354.
- Financial Stability Board (2017). Artificial intelligence and machine learning in financial services.
- Kohlscheen, E., Murcia Pabón, A., and Contreras, J. (2018). Determinants of bank profitability in emerging markets.
- Landauer, T. K., Foltz, P. W., and Laham, D. (1998). An introduction to latent semantic analysis. *Discourse processes*, 25(2-3):259–284.
- Loughran, T. and McDonald, B. (2011). When is a liability not a liability? textual analysis, dictionaries, and 10-ks. *The Journal of finance*, 66(1):35–65.
- Loughran, T. and McDonald, B. (2016). Textual analysis in accounting and finance: A survey. *Journal of Accounting Research*, 54(4):1187–1230.
- Otoritas Jasa Keuangan (2021). Cetak biru transformasi digital perbankan.
- Syamsudin, M. (2022). Stabilitas keuangan dalam menghadapi pandemi covid-19: Efek transformasi sistem keuangan terhadap financial performance perbankan indonesia. *Buletin Riset Kebijakan Perbankan*, 3(2):166–193.
- Yusgiantoro, I. (2018). The adoption of digital banking technology and bank efficiency: Evidence in indonesia. *Working Paper Otoritas Jasa Keuangan*, (18/01).

## A Daftar kata kunci untuk ekstraksi indikator digital awareness dari laporan tahunan

### Kelompok: umum

- automasi
- automation
- budaya digital
- digital
- digital culture
- digital mindset
- digital talent
- digitalisasi
- digitalisation
- digitalization
- inovasi
- innovation
- otomasi
- otomatisasi
- talenta digital
- technology
- teknologi

### Kelompok: apps & development

- agile
- app
- apps
- apple
- google
- play store
- scrum
- user experience
- UX

### Kelompok: digital channel

- e-banking
- ebanking
- internet banking
- mobile
- multi-channel
- multi channel
- omni-channel
- omni channel
- online

### Kelompok: B2B

- API
- application programming
- B2B
- e-commerce
- fintech
- marketplace
- open banking
- platform

- programming interface

### Kelompok: infrastruktur

- cloud
- digital infrastructure
- infrastruktur digital
- investasi digital
- investasi teknologi
- investasi IT

### Kelompok: cyber security

- 3ds
- 3dsecure
- 3d secure
- 3-d secure
- authentication
- cyber
- fraud
- security
- siber

### Kelompok: data

- analytic
- analytics
- big data
- data governance
- data management
- data pribadi
- data processing
- data protection
- data science
- data scientist
- manajemen data
- pengelolaan data
- perlindungan data
- personal data
- predictive

### Kelompok: advanced technology

- artificial intelligence
- AI
- AR
- augmented reality
- biometric
- blockchain
- cardless
- contactless
- credit score
- credit scoring
- digital on-boarding
- digital onboarding
- digital on boarding
- digital signature

- distributed ledger
- kecerdasan buatan
- internet of things
- internet-of-things
- IOT
- machine learning
- ML
- pembelajaran mesin
- pemelajaran mesin
- QR
- QRIS
- regulatory technology
- regtech
- robo
- robot
- robotic
- tanda tangan digital
- tanpa kartu
- video banking
- video on-boarding
- video onboarding
- video on boarding
- VR
- virtual banking
- virtual reality
- virtual assistant

## B Hasil uji signifikansi beberapa model tambahan

Tabel 4: Hasil uji signifikansi dengan metode *robust regression*

	coef	std err	z	P> z	[0.025	0.975]
intercept	-0.0143	0.033	-0.428	0.668	-0.080	0.051
ldr	0.0034	0.004	0.765	0.444	-0.005	0.012
nim	-0.0799	0.090	-0.891	0.373	-0.256	0.096
bopo	0.0131	0.013	1.000	0.317	-0.013	0.039
log(assets)	0.0011	0.003	0.409	0.683	-0.004	0.007
pc1	-0.0013	0.001	-1.216	0.224	-0.003	0.001
pc2	0.0010	0.001	0.721	0.471	-0.002	0.004
pc3	0.0029	0.002	1.785	0.074	-0.000	0.006
pc4	0.0009	0.002	0.540	0.589	-0.002	0.004
pc5	0.0011	0.002	0.608	0.543	-0.002	0.005
pc6	0.0002	0.002	0.096	0.924	-0.004	0.004

Tabel 5: Hasil uji signifikansi dengan metode *clustering* dalam membentuk indikator *digital awareness*

	coef	std err	z	P> z	[0.025	0.975]
intercept	-0.0278	0.032	-0.882	0.378	-0.090	0.034
ldr	0.0033	0.004	0.808	0.419	-0.005	0.011
nim	-0.0440	0.079	-0.554	0.580	-0.200	0.112
bopo	0.0134	0.013	1.064	0.287	-0.011	0.038
log(assets)	0.0004	0.002	0.255	0.798	-0.003	0.003
cl1	0.0250	0.040	0.634	0.526	-0.052	0.102
cl2	0.0014	0.004	0.385	0.700	-0.006	0.008
cl3	0.0012	0.006	0.206	0.837	-0.010	0.013
cl4	-0.0072	0.008	-0.887	0.375	-0.023	0.009