

Perancangan Model Pemetaan Posisi Layanan Bank Jenius Berdasarkan Ulasan Konsumen Menggunakan Teknik *Text Mining*

Diva Awanisa Nahdi
14418033

Dosen Pembimbing:
Prof. Dr. Rajesri Govindaraju, S.T., M.T.

Program Studi Manajemen Rekayasa
Fakultas Teknologi Industri
Institut Teknologi Bandung
Tahun 2022



Daftar Isi

1 Pendahuluan

2 Metodologi

3 Pengolahan Data

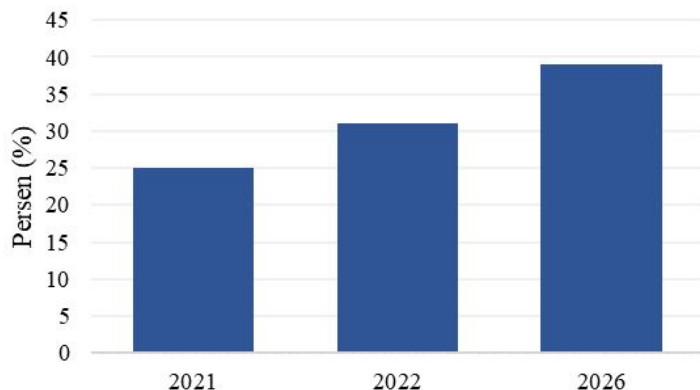
4 Analisis

5 Penutup

Latar Belakang (1)

Potensi pertumbuhan bank digital di Indonesia sangat baik

Proyeksi Persentase Nasabah Bank Digital di Indonesia (2021-2026)



Sumber: Databoks (2021)

Jenius adalah salah satu pemain dalam industri perbankan digital

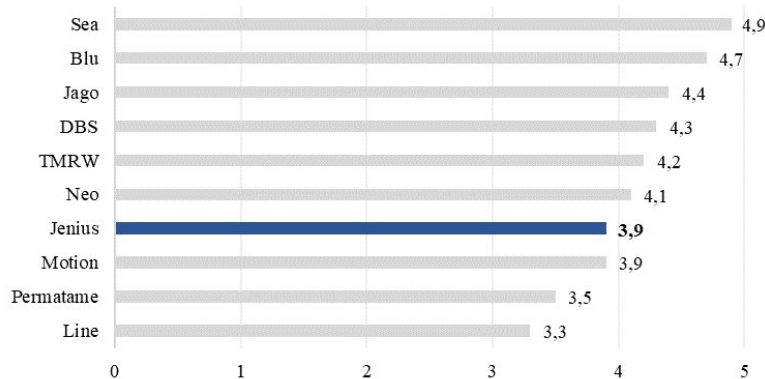


Adanya kompetisi

Latar Belakang (2)

Rating aplikasi Jenius cenderung lebih rendah dari kompetitornya

Rating Aplikasi pada Google Play



Efek rating lebih rendah

- **Purchase intention** menurun
(Santosa & Jatra, 2022)
- Konsumen cenderung memilih aplikasi **rating > 4.0**
(Banks.com dan Northwestern University, 2016)
- Menunjukkan adanya ketidakpuasan konsumen eksisting → **meningkatkan churn rate**
(Roberts dkk., 2021)

Terdapat **urgensi** untuk menyelesaikan masalah ini

Latar Belakang (3)



Latar Belakang (4)



Rumusan Masalah

Bagaimana model **pemetaan posisi** layanan Jenius menurut **ulasan konsumen** menggunakan teknik *text mining*?"

Pemangku Kepentingan



Problem Owner
Product Manager Jenius



Problem User
Tim produk Jenius



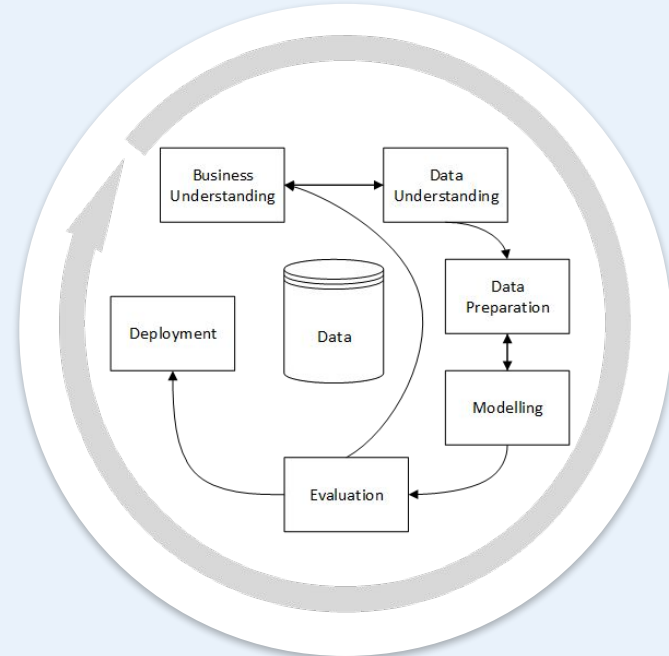
Problem Customer
Nasabah Jenius



Problem Analyst
Peneliti

Standar Acuan (1)

Cross-Industry Standard Procedure for Data Mining (CRISP-DM)



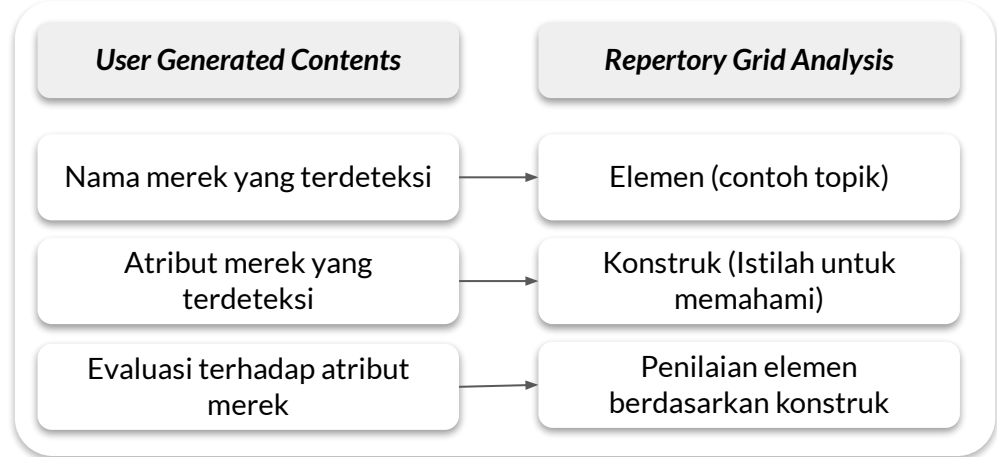
Standar Acuan (2)

Pemodelan Pemetaan Posisi oleh Hu dan Trivedi (2022)

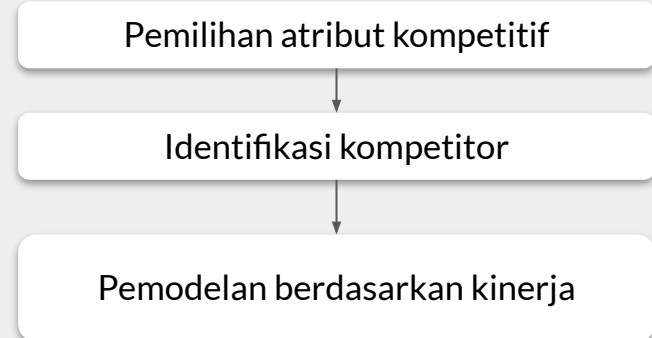
Memetakan merek menggunakan pendekatan *Repertory Grid Analysis*

Metode yang digunakan untuk memahami bagaimana seseorang dapat **menginterpretasikan** sesuatu dalam suatu **konstruk**

Ulasan konsumen sebagai *input* untuk melakukan RGA



Tahapan



Rumusan Masalah

1. **Atribut** apa saja yang dipertimbangkan dalam **memetakan** layanan bank digital?
2. Bagaimana **kompetitor utama** bank digital ditentukan berdasarkan atribut-atribut yang telah diidentifikasi?
3. Bagaimana **perbandingan kinerja** bank digital terhadap kompetitor utama berdasarkan atribut yang telah diidentifikasi?

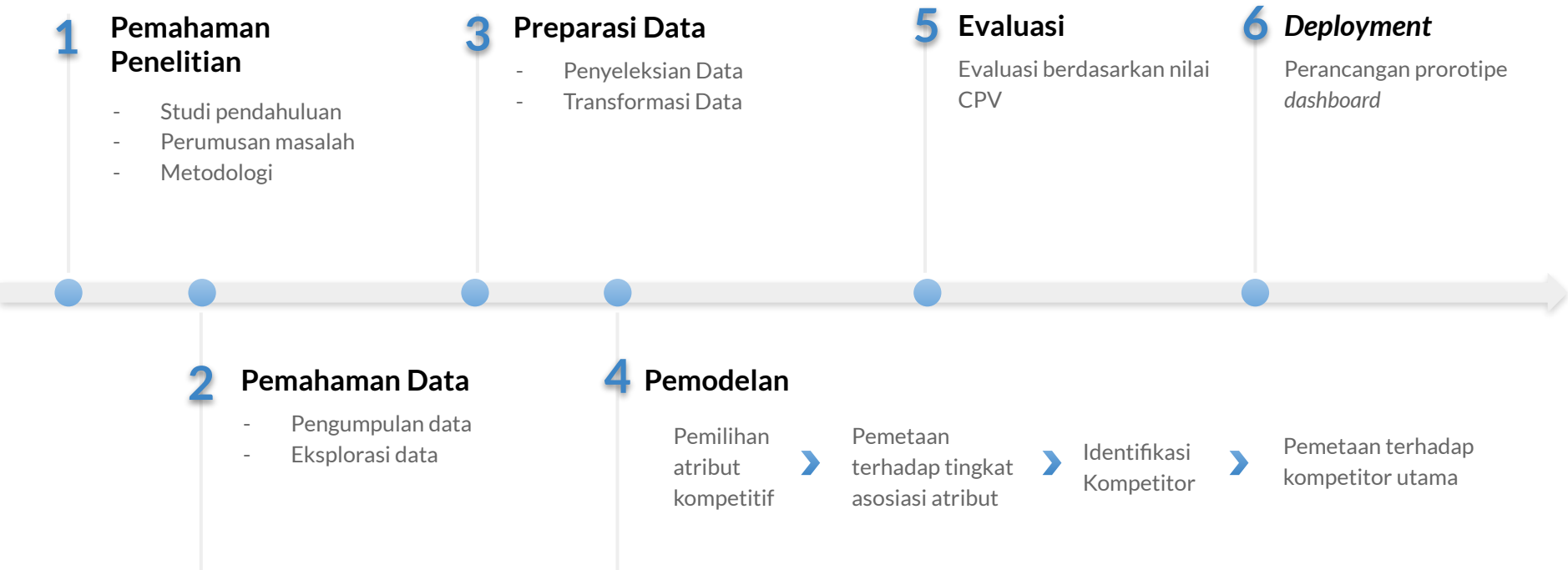
Tujuan

1. Menentukan **atribut yang paling diasosiasikan** oleh konsumen terhadap merek
2. Menentukan **kompetitor utama** bank digital
3. Merancang **model pemetaan** posisi Jenius terhadap **kompetitor utama**

Manfaat

1. Memberikan **gambaran persepsi konsumen** terhadap merek bank digital
2. Memberikan **tingkat prioritas** dari pengembangan atau perbaikan atribut layanan bank digital

Berdasarkan *framework Cross Industry Standard Process for Data Mining (CRISP-DM)*



Pengumpulan dan Eksplorasi Data

Dilakukan *web scrapping* pada **Google Play** menggunakan Python terhadap 7 aplikasi

**Jumlah Data**

| Merek | Jumlah |
|----------|--------|
| Neobank | 143480 |
| Jenius | 113371 |
| Seabank | 35897 |
| Jago | 27974 |
| Digibank | 27313 |
| TMRW | 14335 |
| Blu | 13441 |

Kolom Data

| Nama kolom | Type data |
|----------------------|-----------|
| reviewId | String |
| username | String |
| userImage | String |
| Content | String |
| Score | Integer |
| thumbsUpCount | Integer |
| reviewCreatedVersion | String |
| At | String |
| replyContent | String |
| repliedAt | String |

Penyeleksian Data

Menyeleksi data yang relevan

Penyaringan berdasarkan waktu

Menyamarkan titik awal proses pengolahan data

→ menghindari adanya pengaruh **tren** yang **berbeda**

1 Juli 2021

26 Agustus 2022

Seluruh aplikasi kompetitor sudah diluncurkan

Waktu pengumpulan data

Pemotongan data

- Menyesuaikan kapasitas memori perangkat keras
- Maksimal 20.000, dipilih secara acak

| Merek | Sebelum | Sesudah |
|---------|---------|---------|
| Neobank | 130648 | 20000 |
| Jenius | 39161 | 20000 |
| Seabank | 35792 | 20000 |
| Jago | 26202 | 20000 |

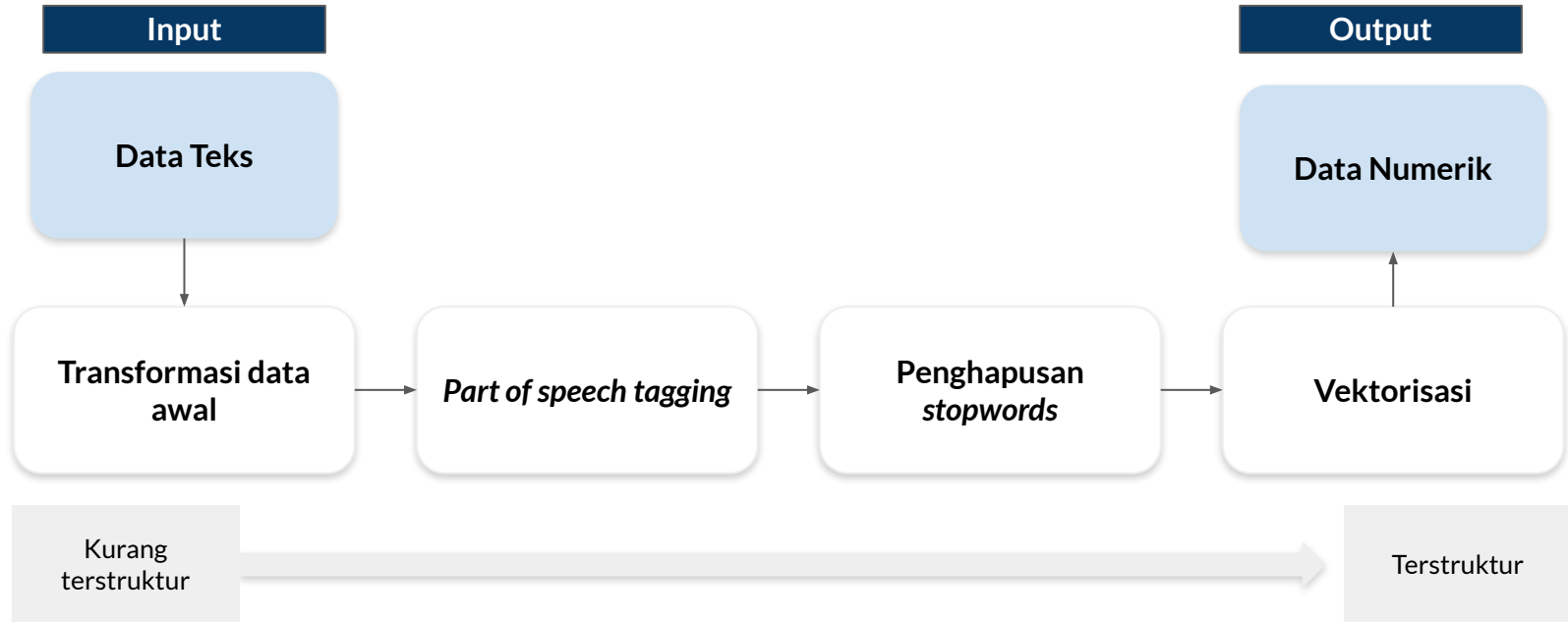
Seleksi kolom

Memilih kolom yang relevan

| | |
|---------|----------------------|
| Content | Ulasan yang diunggah |
| Score | Nilai <i>rating</i> |

Transformasi Data (1)

Mengubah bentuk data tidak terstruktur → terstruktur



Transformasi Data (2)

Mengubah bentuk data tidak terstruktur → terstruktur

1

Pengecilan Huruf

- Untuk menyeragamkan bentuk kata
- Menghindari deteksi dua kata yang berbeda

| Sebelum | Sesudah |
|--|--|
| Pindah device aja ribet banget mesti tlp vs segala. Padahal bukan ganti muka atau ganti nomer. Malesin banget jago Sekarang!!! | pindah device aja ribet banget mesti tlp vs segala. padahal bukan ganti muka atau ganti nomer. malesin banget jago sekarang!!! |

2

Penghapusan angka, *whitespace* dan emoji

- Untuk menyeragamkan bentuk kata
- Sentimen sudah diwakilkan oleh nilai *rating* pada setiap ulasan

| Sebelum | Sesudah |
|--|---|
| bagus 🍌🍌🍌🍌🍌 | bagus |
| saya ada keluhan .. tranfer 3jt udah 2 hari tidak terkirim. mintadi cancel tidak di respon | saya ada keluhan .. tranfer jt udah hari tidak terkirim. mintadi cancel tidak di respon |
| mantap | mantap |

Transformasi Data (3)

Mengubah bentuk data tidak terstruktur → terstruktur

3

Perbaikan kata

Agar makna kata dapat terbaca oleh model *pos tagger*

| Sebelum | Sesudah |
|---------|---------|
| tp | tetapi |
| byk | banyak |

4

Perbaikan Huruf Berulang

- Agar makna kata dapat terbaca oleh model *pos tagger*
- Memperbaiki huruf yang terhapus

| Sebelum | Sesudah |
|-----------------------------------|-----------------------------------|
| jago syariah kerennnnn | jago syariah keren |
| aplikasi mudah dan banyak manfaat | aplikasi mudah dan banyak manfaat |

Transformasi Data (4)

Mengubah bentuk data tidak terstruktur → terstruktur

5

Penyambungan kata

- Terbatas pada se- dan nya-

| Sebelum | Sesudah |
|--------------------------------------|-------------------------------------|
| bank jago, dunia baru dunia nya jago | bank jago, dunia baru dunianya jago |

6

Pemisahan kata

- Agar dapat diproses pada model *part-of-speech tagging*
- Pemisahan dilakukan berdasarkan spasi dan tanda baca

| Sebelum | Sesudah |
|--------------------------------------|--|
| terima kasih banyak. sangat membantu | ['terima', 'kasih', 'banyak', '.', 'sangat', 'membantu'] |

Transformasi Data (5)

Mengubah bentuk data tidak terstruktur → terstruktur

7

Part-of-speech tagging

- Melabelkan kata berdasarkan kelompok katanya
- Hanya mengambil **kata benda/ noun**

| Sebelum | Sesudah <i>Part-of-speech tagging</i> | Sesudah Penyaringan Kelompok Kata |
|--|---|-----------------------------------|
| ['aplikasi', 'tidak', 'jelas', 'susah', 'login'] | [('aplikasi', 'NN'), ('tidak', 'NEG'), ('jelas', 'JJ'), ('susah', 'VB'), ('login', 'RB')] | ['aplikasi'] |

8

Penghapusan stopwords

- Menghapus kata yang tidak mengandung informasi
- Menyaring hasil *pos tagging* agar lebih bersih

| Sebelum | Sesudah |
|--|---|
| ['aplikasi', 'lupa', 'password', 'malah', 'verifikasi', 'langkah', 'cs'] | ['aplikasi', 'lupa', 'password', 'verifikasi', 'langkah', 'cs'] |

Transformasi Data (6)

Mengubah bentuk data tidak terstruktur → terstruktur

9

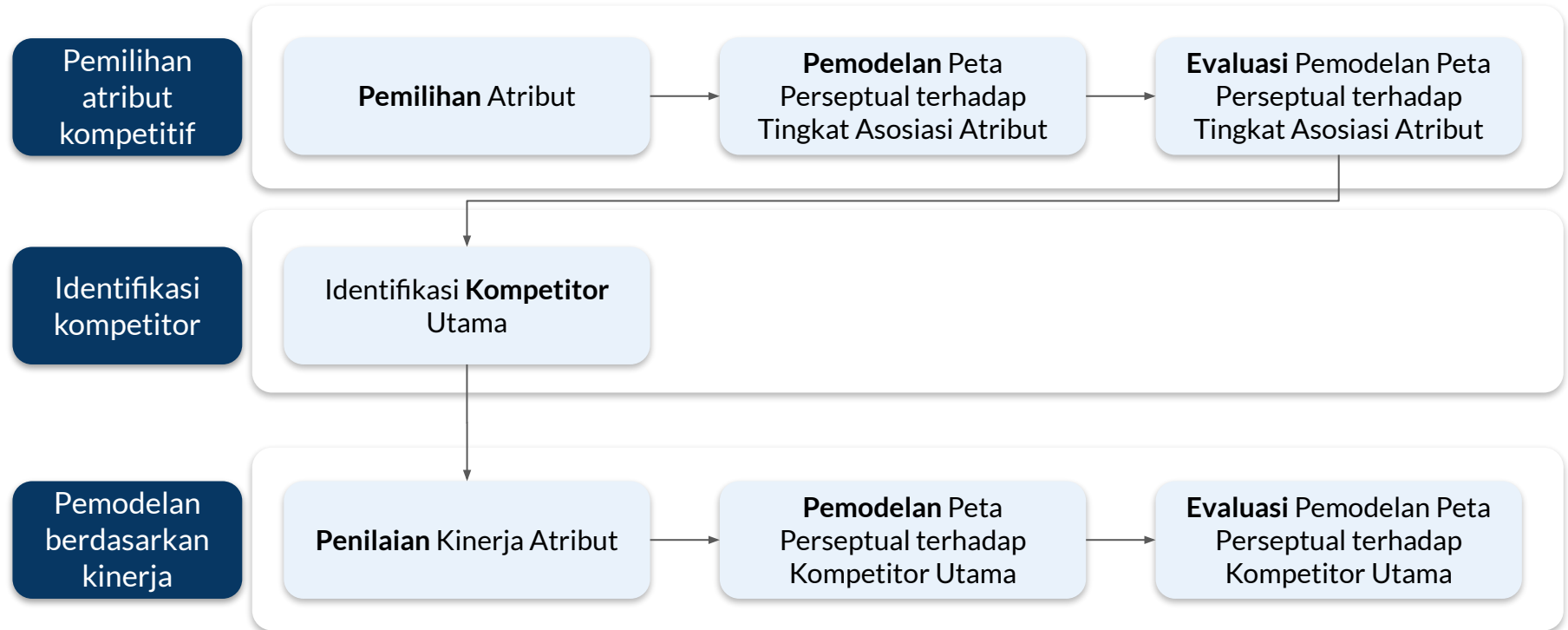
Vektorisasi data

- Bentuk terstruktur dari data
- 1 -2 kata
- Berdasarkan ada (1) atau tidak ada (0)

Ilustrasi

| | bank | bank jago | jago | jago keren | jago mantap | keren | mantap |
|------------------|------|-----------|------|------------|-------------|-------|--------|
| bank jago mantap | 1 | 1 | 1 | 0 | 1 | 0 | 1 |
| bank jago keren | 1 | 1 | 1 | 1 | 0 | 1 | 0 |

Pemodelan



Pemilihan Atribut

Untuk menentukan **atribut yang kompetitif** dengan merek lain.
Dipilih berdasarkan tingkat **asosiasi** atribut terhadap merek

Jaccard Similarity

$$\frac{a}{F1 + F2 - a}$$

a : Frekuensi dokumen
dengan atribut pada merek

F1: Frekuensi dokumen
pada merek

F2: Frekuensi dokumen
dengan atribut pada semua
merek

Koefisien Jaccard

Pemilihan dilakukan
berdasarkan **Koefisien
Jaccard tertinggi**

Dipilih **30 atribut** yang
menggambarkan
kompetisi industri
perbankan digital

Atribut terpilih

| | | |
|----------|------------|------------|
| Akun | Hp | Pin |
| Biaya | Isi | Promo |
| Cs | Jam | Proses |
| Daftar | Kartu | Rekening |
| Dana | Kode | Tabungan |
| Data | Login | Tampilan |
| Download | Nasabah | Tarik |
| Email | Nomor | Transaksi |
| Error | Pelayanan | Uang |
| Fitur | Pembayaran | Verifikasi |

Pemodelan Terhadap Tingkat Asosiasi Atribut (1)

Menggunakan *Principal Component Analysis*

Atribut Terpilih

Pemeringkatan

Normalisasi

Pemilihan Dimensi

PCA

Untuk menghilangkan efek ketidakseimbangan data

Salah satu tahapan pemodelan PCA

Secara *a priori*
→ **3 Dimensi**

- *Factor Loading*
- *Element Loading*
- *CPV*

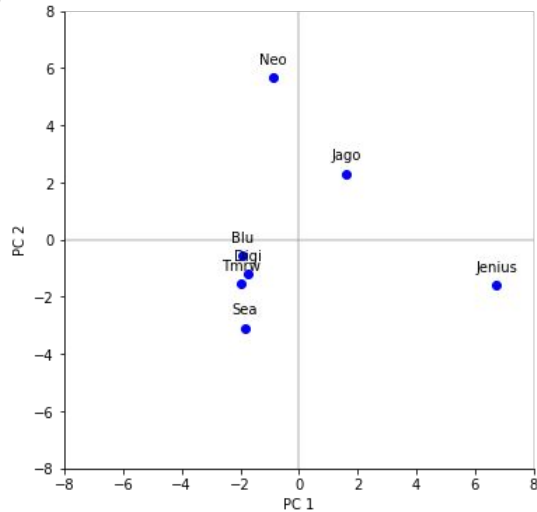
| PC 1 | | PC 2 | | PC 3 | |
|------------|---------|-----------|---------|-------|---------|
| Rekening | 0.87 | Proses | 0.9232 | Data | -0.8232 |
| Hp | -0.8119 | Kode | -0.8252 | Isi | 0.7599 |
| Pembayaran | -0.7953 | Transaksi | 0.8135 | Error | -0.7484 |
| Pin | -0.7788 | Cs | 0.734 | Login | -0.6248 |
| Tampilan | -0.7229 | Jam | -0.7071 | Email | -0.6113 |
| Verifikasi | 0.7111 | Biaya | 0.6542 | Uang | 0.4858 |
| Download | -0.7093 | Kartu | 0.6226 | Tarik | 0.4435 |
| Daftar | 0.6587 | Pelayanan | 0.6177 | | |
| Nomor | 0.612 | Fitur | 0.5325 | | |
| Dana | 0.5925 | | | | |
| Nasabah | 0.5365 | | | | |
| Tabungan | 0.4398 | | | | |
| Promo | 0.4068 | | | | |
| Akun | -0.378 | | | | |

Factor Loading

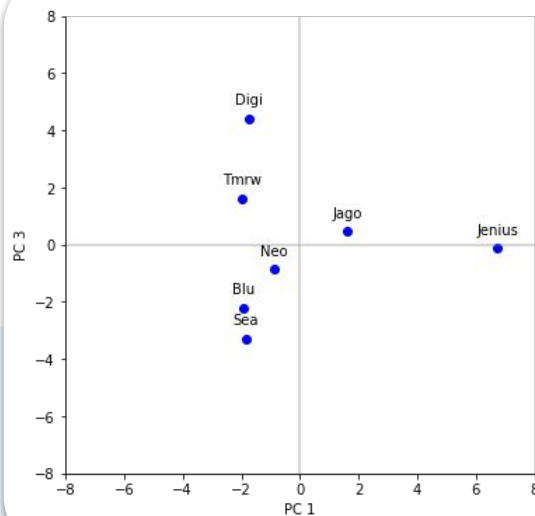
Pemodelan Terhadap Tingkat Asosiasi Atribut (2)

Pemetaan dilakukan terhadap *element loading* yang dihasilkan oleh pemodelan

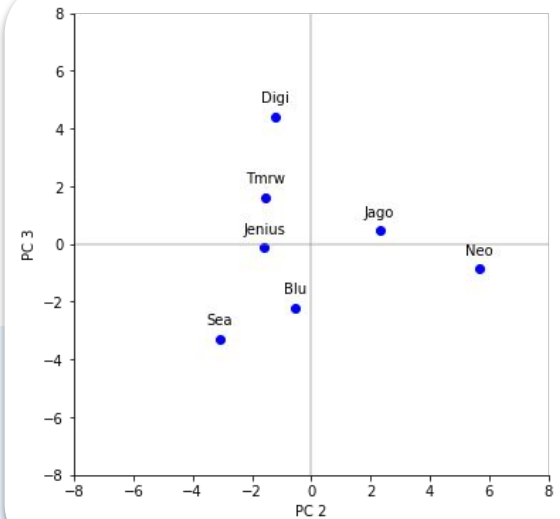
PC 1 vs PC 2



PC 1 vs PC 3



PC 2 vs PC 3



Evaluasi Pemodelan Terhadap Tingkat Asosiasi Atribut

Cumulative Percentage Variance

Proporsi variansi yang menjelaskan dalam model

Batas minimum: **50%** (Alzate, 2022)



CPV Model

74 %

Performa model sudah **baik**

Identifikasi Kompetitor Utama

Untuk memfokuskan perumusan strategi

Metode Slater

Berdasarkan jarak *Euclidean* yang dinormalisasi

Normalisasi

Untuk **menghilangkan efek *outlier***, sehingga setiap merek tetap memiliki kompetitor utama

Jarak *Euclidean*

Panjang hipotenusa titik pada setiap dimensi. Pengukuran berdasarkan besar ***Element Loading***

Kompetitor

| Merek | Kompetitor 1 | Kompetitor 2 |
|---------------|--------------|--------------|
| Blu | Sea | Tmrw |
| Digi | Jago | Tmrw |
| Jago | Blu | Neo |
| Jenius | Jago | Tmrw |
| Neo | Blu | Jago |
| Sea | Blu | Tmrw |
| Tmrw | Blu | Digi |

Penilaian Kinerja Atribut

Sebagai acuan dalam menentukan arah pengembangan strategi

Binomial Proportion Test

Untuk melihat distribusi sentimen konsumen terhadap atribut



Nilai Z

- **Positif:**
Semakin besar, performa semakin baik
- **Negatif:**
Semakin besar, performa semakin tidak baik

Nilai Z Atribut

| Atribut | Jenius | Jago | Tmrw |
|------------|--------|--------|--------|
| Akun | -15.92 | -22.64 | -10.29 |
| Biaya | -7.17 | 6.11 | 11.69 |
| Cs | -20.2 | -14.39 | -7.62 |
| Daftar | -12.03 | -17.67 | -12.57 |
| ... | ... | ... | ... |
| Verifikasi | -13.36 | -19.17 | -18.25 |

Pemodelan terhadap Kompetitor Utama (1)

Menggunakan *Principal Component Analysis*

Performa Atribut

Normalisasi

Pemilihan Dimensi

PCA

Salah satu tahapan pemodelan PCA

Secara *a priori*
→ **2 Dimensi**

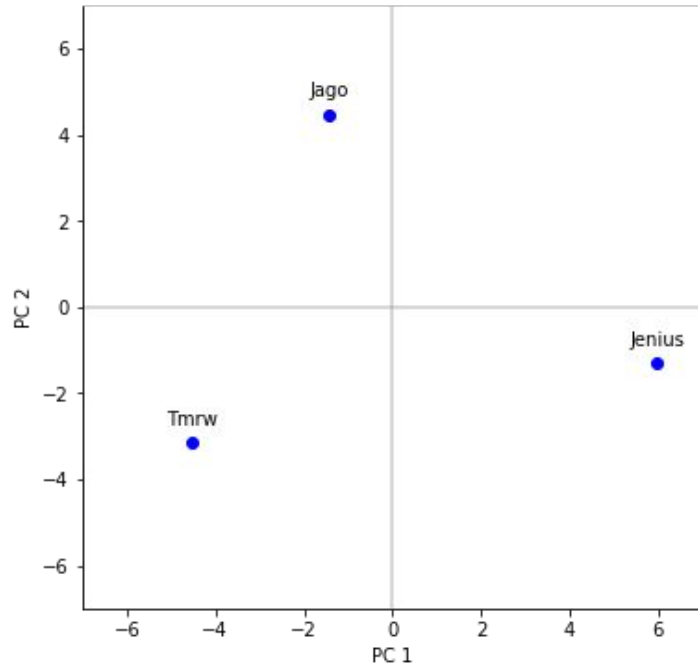
- *Factor Loading*
- *Element Loading*
- *CPV*

Factor Loading

| PC 1 | | PC 2 | |
|-----------|---------|--------|---------|
| Biaya | -0.9798 | Akun | -0.9516 |
| Rekening | 0.9724 | Dana | -0.9484 |
| Login | -0.972 | Daftar | -0.9303 |
| Tabungan | 0.9689 | Nomor | -0.9155 |
| Proses | 0.9654 | Kartu | 0.9062 |
| Fitur | 0.9616 | Promo | -0.8528 |
| Pin | -0.9576 | Data | -0.8196 |
| Tampilan | 0.9438 | Eror | 0.7921 |
| Cs | -0.9426 | Email | -0.7178 |
| Pelayanan | -0.9339 | Jam | -0.7158 |
| Tarik | -0.9317 | Uang | -0.7089 |
| ... | ... | | |

Pemodelan terhadap Kompetitor Utama (2)

Pemetaan dilakukan terhadap *element loading* yang dihasilkan oleh pemodelan



| PC 1 | |
|-------|---------|
| Biaya | -0.9798 |
| Login | -0.972 |
| Pin | -0.9576 |
| ... | ... |

| PC 1 | |
|----------|--------|
| Rekening | 0.9724 |
| Tabungan | 0.9689 |
| Proses | 0.9654 |
| ... | ... |

| PC 2 | |
|--------|---------|
| Akun | -0.9516 |
| Dana | -0.9484 |
| Daftar | -0.9303 |
| ... | ... |

| PC 2 | |
|-------|--------|
| Kartu | 0.9062 |
| Error | 0.7921 |

Evaluasi Pemodelan terhadap Kompetitor Utama

Cumulative Percentage Variance

Proporsi variansi yang ter jelaskan dalam model

Batas minimum: 50%



CPV Model

100 %

Performa model sudah **baik**

Perancangan *Dashboard* (1)

Untuk menyebarkan hasil pemodelan

Dash

Package untuk membuat *Dashboard* yang interaktif

Heroku

Server yang dapat menyimpan hasil pemodelan dan prototipe
<https://komp-jenius.herokuapp.com/>

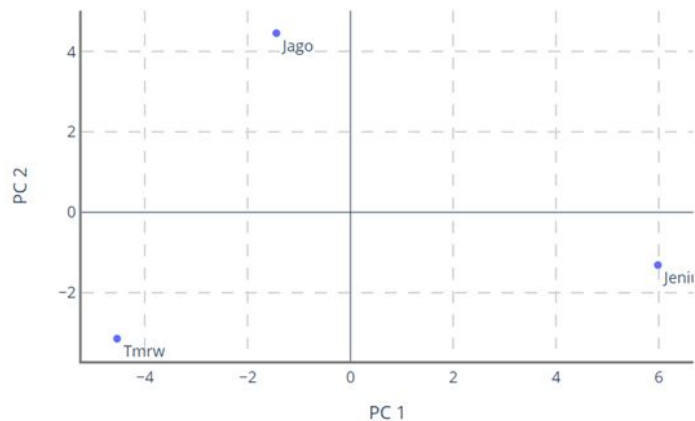
DASHBOARD PETA PERSEPTUAL

| Atribut PC 2 | Korelasi |
|--------------|----------|
| Kartu | 0.906 |
| Error | 0.792 |



PETA

| Atribut PC 1 | Korelasi |
|--------------|----------|
| Blaya | -0.98 |
| Login | -0.972 |
| Pin | -0.958 |
| Cs | -0.943 |



| Atribut PC 1 | Korelasi |
|--------------|----------|
| Rekening | 0.972 |
| Tabungan | 0.969 |
| Proses | 0.965 |
| Fitur | 0.962 |

| Atribut PC 2 | Korelasi |
|--------------|----------|
| Akun | -0.952 |
| Dana | -0.948 |
| Daftar | -0.93 |
| Nomor | -0.915 |

Perancangan *Dashboard* (2)

Untuk menyebarkan hasil pemodelan



Fitur zoom in, zoom out, drag, select

Analisis Validasi

Melakukan wawancara dengan *problem user*

| Proses | Tanggapan | Keterangan |
|------------------------------|---|------------|
| Pemilihan atribut | Sudah menggambarkan kondisi pasar | Valid |
| Pemetaan terhadap kompetitor | Perbandingan performa terhadap kompetitor utama sudah benar | Valid |
| Perancangan <i>Dashboard</i> | Peletakan grafik dan tabel sudah mudah diinterpretasi | Valid |

Kesimpulan

Atribut apa saja yang dipertimbangkan dalam memetakan layanan bank digital?



Atribut yang paling diasosiasikan oleh konsumen terhadap merek ditentukan berdasarkan **Koefisien Jaccard**.

'akun', 'biaya', 'cs', 'daftar', 'dana', 'data', 'download', 'email', 'eror', 'fitur', 'hp', 'isi', 'jam', 'kartu', 'kode', 'login', 'nasabah', 'nomor', 'pelayanan', 'pembayaran', 'pin', 'promo', 'proses', 'rekening', 'tabungan', 'tampilan', 'tarik', 'transaksi', 'uang', dan 'verifikasi'

Bagaimana kompetitor utama bank digital ditentukan berdasarkan atribut-atribut yang telah diidentifikasi?



Kompetitor utama ditentukan menggunakan Metode Slater. Kompetitor utama dari **Bank Jenius** adalah **Bank Jago** dan **TMRW**

Bagaimana **perbandingan kinerja** atribut bank digital Jenius terhadap kompetitor utama?



Jenius **unggul** dalam atribut **rekening, tabungan, dan proses** dan **tidak unggul** dalam atribut **biaya, login, dan pin**

- **TMRW: baik pada eror dan kartu, buruk pada akun, dana, pendaftaran**
- **Jago: baik pada akun, dana, pendaftaran, buruk pada eror dan kartu**

Saran

Untuk Perusahaan

1. Merumuskan strategi kompetitif dengan **memprioritaskan** atribut layanan yang **memiliki nilai korelasi paling tinggi** pada masing-masing dimensi
2. Membuat **sistem pengoleksian data** yang terintegrasi dengan prototipe *dashboard* yang telah dibuat dan mampu menampilkan hasil model dengan **periode waktu yang fleksibel**
3. Melakukan diskusi untuk **menginterpretasikan konteks** dari atribut.

Untuk Penelitian Selanjutnya

1. Melakukan **pemodelan** yang mampu **menggambarkan konteks** dari atribut yang dimodelkan
2. Melakukan **hyperparameter tuning** dengan membandingkan hasil dari pemodelan PCA secara linear dengan hasil pemodelan PCA lain
3. Melakukan **proses stemming** untuk menstandarisasikan kata yang bermakna sama
4. Melakukan **proses translasi** dari bahasa asing untuk menyeragamkan konteks dari kata yang diproses

Terima Kasih.

Asumsi

1. Data yang diolah dapat **merepresentasikan populasi** konsumen bank digital
2. Ulasan yang diolah **bukan ulasan palsu**
3. Pendapat dari konsumen **tidak berubah** sejak periode pengumpulan data sampai pembuatan model.
4. Sentimen dari setiap kalimat dalam ulasan konsumen dapat **direpresentasikan oleh rating** keseluruhan yang diberikan oleh konsumen

Batasan

1. Data yang digunakan adalah data dari tanggal **1 Juli 2021** sampai **26 Agustus 2022**
2. Merek bank digital yang diolah:
 - Bank yang paling populer
 - Nilai *rating* lebih tinggi dari Jenius

Analisis Keterbatasan

1

Hanya memetakan **6 kompetitor**

2

Atribut **tidak** memiliki **konteks**
yang lengkap

3

Perangkat keras memiliki
spesifikasi yang terbatas

4

Tidak meninjau **ulasan palsu**

5

Pemilihan atribut masih
dilakukan **manual**

6

Tidak melibatkan **stemming**

Implikasi Manajerial

1

Sebaiknya memprioritaskan **perbaikan atau pengembangan** atribut yang **berkorelasi tinggi** dengan masing-masing dimensi

2

Membuat sistem **pengoleksian data**

3

Mengembangkan **dashboard** yang mencakup **penyeleksian** berdasarkan **periode waktu**

4

Menggunakan perangkat keras dengan **spesifikasi lebih baik**

5

Melakukan diskusi untuk **menginterpretasi konteks** atribut