

## PENERAPAN METODE K-MEANS UNTUK SEGMENTASI KELUHAN PELANGGAN (study kasus : Hanatex)

Nur Ika Royanti<sup>1)</sup>, Indrayanti<sup>2)</sup>, Nurul Amalia<sup>3)</sup>  
STMIK Widya Pratama <sup>123)</sup>

ikaroyant@gmail.com<sup>1)</sup>, indrayanti3214@gmail.com<sup>2)</sup>, amalia.0121@gmail.com<sup>3)</sup>

### Abstrak

*Benang merupakan bahan baku utama dalam pembuatan produk tekstil. Banyaknya pengrajin tekstil di Indonesia khususnya Pekalongan mendorong pengusaha produksi benang harus terus berjalan. Dengan banyaknya pesaing di industri tekstil ini pihak penjual semakin sulit untuk mendapat kepercayaan dari pihak konsumen. Selain melihat dari sisi harga, konsumen juga akan menilai dari kualitas produk yang diterimanya. Hal ini perlu diperhatikan untuk mempertahankan loyalitas konsumen terhadap pembelian ulang produk. Salah satu pengrajin benang di pekalongan yaitu Hanatex. Meskipun transaksi pembelian benang terus berjalan, akan tetapi Hanatex masih sering menerima keluhan dari konsumen terhadap kualitas benang yang diterimanya. Keluhan yang di didapatkan dari konsumen bisa menjadi kesempatan untuk memperbaiki kesalahan sekaligus meningkatkan kualitas bisnis. Hal tersebut perlu di perhatikan dan ditangani dengan benar agar nantinya dapat menjadi bahan pertimbangan dalam pengambilan keputusan. Pemenuhan kualitas layanan yang memuaskan konsumen salah satunya yaitu dengan merespon keluhan dari konsumen. Keluhan pelanggan yang bervariasi perlu dilakukan segmentasi untuk mengelompokkan data-data berdasarkan keluhan tertentu yang lebih spesifik. Dengan demikian dapat memudahkan dalam melakukan analisa keluhan pelanggan untuk meningkatkan pelayanan dengan perbaikan yang tepat. Berdasarkan hasil evaluasi menggunakan DBI performance dari algoritma k-means dalam melakukan segmentasi dapat disimpulkan baik karena 88,8% kemiripan data dalam kelompok mirip. Untuk penelitian berikutnya disarankan perlu ditambahkan metode untuk cleaning data pada text processing untuk menghilangkan kata singkatan dan kata tidak baku agar dapat meningkatkan performance dan akurasi metode.*

**Kata kunci:** K-Means, Segmentasi, Keluhan Pelanggan

### 1. Pendahuluan

Salah satu industri yang mempunyai kontribusi besar pada perekonomian di Indonesia yaitu tekstil dan produk tekstil (TPT) (Riyanto 2020). Produksi tekstil di Indonesia telah terintegrasi dari mulai pembuatan benang, rajutan hingga menjadi produk akhir siap pakai. Benang merupakan bahan baku utama dalam pembuatan produk tekstil sehingga benang menjadi salah satu produksi yang menjadi perhatian khusus (Kusumo and Riandadari 2017). Banyaknya pengrajin tekstil di Indonesia khususnya Pekalongan mendorong pengusaha produksi benang harus terus berjalan. Dengan banyaknya pesaing di industri tekstil ini pihak pengrajin atau penjual semakin sulit untuk mendapat

kepercayaan dari pihak konsumen (Amalia, Royanti, and Ismanto 2023).

Cara pandang konsumen terhadap barang berubah seiring berkembangnya teknologi informasi. Selain melihat dari sisi harga, konsumen juga akan menilai dari kualitas produk yang diterimanya. Hal ini perlu diperhatikan untuk mempertahankan loyalitas konsumen terhadap pembelian ulang produk (Napitupulu 2019).

Salah satu pengrajin benang di pekalongan yaitu Hanatex. Meskipun transaksi pembelian benang terus berjalan, akan tetapi Hana Text masih sering menerima keluhan dari konsumen terhadap kualitas benang yang diterimanya. Keluhan yang di didapatkan dari konsumen bisa menjadi kesempatan untuk

memperbaiki kesalahan sekaligus meningkatkan kualitas bisnis (Al'asqolaini 2019). Keluhan dari konsumen masing-masing berbeda tergantung dari kekurangan dari produk yang di terima dan disampaikan lewat pesan whatsapp. Hal tersebut perlu di perhatikan dan ditangani dengan benar agar nantinya dapat menjadi bahan pertimbangan dalam pengambilan keputusan.

Dalam mempertahankan loyalitas konsumen dalam pembelian barang, terdapat beberapa faktor yang berpengaruh diantaranya kualitas layanan dan kualitas produk (Putro 2014). Pemenuhan kualitas layanan yang memuaskan konsumen salah satunya yaitu dengan merespon keluhan dari konsumen. Untuk menjaga kualitas pelayanan dan tetap menjaga hubungan yang baik antara perusahaan dan konsumennya dapat melalui penanganan keluhan yang di sampaikan oleh konsumen kepada perusahaan (Ani and Jatmiko 2023). Dengan penanganan keluhan yang efektif juga akan menambah kepuasan konsumen terhadap pelayan yang diberikan pihak perusahaan. Sedangkan, Penanganan keluhan pelanggan yang tidak tepat akan menjadi bumerang bagi perusahaan sendiri karena pelanggan yang kecewa akan dengan cepat menyebarkan kekecewaannya, baik melalui mulut ke mulut (*words of mouth*) ataupun melalui media, baik cetak maupun online (Ashma Nurmeila, Witanti, and Sabrina Nurul 2020) Keluhan pelanggan yang bervariasi perlu dilakukan segmentasi untuk mengelompokkan data-data berdasarkan keluhan tertentu yang lebih spesifik. Dengan demikian dapat memudahkan dalam melakukan analisa keluhan pelanggan untuk meningkatkan pelayanan dengan perbaikan yang tepat.

Beberapa penelitian keluhan pelanggan telah dilakukan diantaranya yaitu menerapkan algoritma k-means untuk melakukan segmentasi pelanggan berdasarkan keluhan diketahui bahwa dalam penelitiannya keluhan pelanggan terbagi menjadi 3 segmentasi penanganan keluhan yaitu sulit, normal dan mudah waktu penanganan keluhan (Ashma Nurmeila, Witanti, and Sabrina Nurul 2020), Analisis penyampaian keluhan konsumen menggunakan metode clustering K-Means yang dilakukan terhadap konsumen pengguna provider Indosat dengan menggunakan data tweet sejumlah 300 data acak yang di

kumpulkan dari bulan desember 2018 hingga bulan april 2019 diperoleh hasil bahwa ada 3 berbentuk wordcloud berbeda, yaitu pada wordcloud cluster 0 anggotanya berbicara tentang jaringan Indosat yang parah, pada wordcloud cluster 1 anggotanya berbicara tentang permintaan perbaikan jaringan sinyal Indosat, dan pada wordcloud cluster 2 anggotanya berbicara tentang jaringan sinyal parah Indosat pada daerah Bogor (Saputra and Arianty 2019).

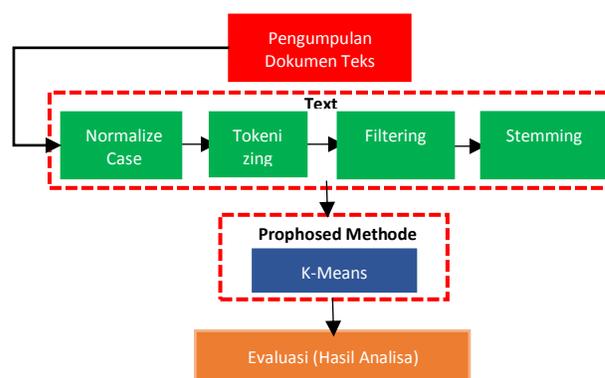
Pada penelitian ini akan dilakukan pengelompokan dengan *text mining* dari data keluhan pelanggan yang ada pada PT Hana Text dengan menggunakan metode K-Means. Tujuannya yaitu untuk menganalisa segmentasi keluhan pelanggan dan dapat mengetahui tingkat ketepatan metode K-Means untuk segmentasi data.

## 2. Metode Penelitian

Setelah menggunakan produk, pelanggan yang puas akan memberikan ulasan yang positif dan akan merekomendasikannya ke orang lain. Sementara, pelanggan yang tidak puas dan merasa kecewa terhadap produk akan mereka kecewa dan memberikan keluhan. keluhan dari pelanggan pada akhirnya dapat menjadi peluang untuk meningkatkan kualitas bisnis apabila ditangani dengan tepat.

### 2.1 Alur Penelitian

Agar penelitian ini dapat berjalan dengan baik, maka perlu dibuatkan alur penelitian. Adapun alur penelitian pada penelitian ini seperti pada Gambar 2.1 yang meliputi:



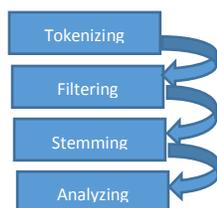
Gambar 1 Alur Penelitian

## 2.2 Pengumpulan Dokumen Text

Data yang digunakan pada penelitian ini diperoleh dari PT Hanatex mengenai keluhan pelanggan yang disampaikan melalui whatsapp, data yang digunakan sebanyak 100 dokumen. Data berupa dokumen text berupa komentar atau keluhan-keluhan pelanggan yang disampaikan mengenai produk benang yang dijual oleh Hanatex. Komentar tersebut akan dilakukan segmentasi berdasarkan keluhan.

## 2.3 Text Processing

Penelitian ini menggunakan data dari Hana Text dengan proses pengolahan text mining yang merupakan suatu proses untuk menggali informasi dari sekumpulan dokumen menggunakan tools analisis yang merupakan komponen-komponen dalam *data mining* (Ariyanti and Iswardani 2020). Data yang akan diolah berupa text sehingga pengolahan data menggunakan *text processing*. Setelah dokumen teks telah diperoleh, tahap selanjutnya adalah text preprocessing yang meliputi:



Gambar 2 Proses *Text Mining*

1. Normalize Case: memproses data teks keluhan yang diperoleh untuk merubah semua bentuk huruf besar menjadi huruf kecil.
2. Tokenizing: melakukan pemecahan dokumen berdasarkan tiap kata
3. Filtering: penghapusan kata-kata yang tidak penting yang tidak memiliki arti dan penghapusan pada kata-kata yang memiliki nilai kecil pada komentar.
4. Stemming: proses pengubahan penggalan kata dasar yang diperoleh dari data komentar yang telah masuk tahap filtering dan Stopword menjadi bentuk kata baku bahasa Indonesia yang baik dan benar

Dokumen text yang sudah siap akan dilakukan pengelompokan/kluster dengan

menggunakan algoritma K-means. Hasil dari proses clusterisasi menggunakan algoritma akan dilakukan pengujian dengan DBI (Davies Bouldin Index) untuk mengetahui performance algoritma k-means dalam mengelompokkan data berdasarkan kemiripan data dalam kelompok. Hasil dari proses tersebut dapat dilakukan analisa untuk mengetahui pengelompokkan dari ulasan yang diberikan pelanggan berdasarkan keluhan pelanggan yang telah disampaikan.

## 3. Hasil dan Pembahasan

Penelitian ini dilakukan dengan memanfaatkan tool RapidMiner 9.10. dalam proses pengolahan data keluhan pelanggan menggunakan algoritma k-means untuk dapat mengetahui segmen-segmen yang menjadi keluhan pelanggan. Adanya model yang terbentuk dari hasil penelitian ini nantinya dapat memudahkan pihak PT Hanatex untuk melakukan perbaikan secara tepat sesuai dengan keluhan dari pelanggan. Beberapa tahapan yang dilakukan dalam penelitian sebagai berikut:

### 3.1 Analisa Data

Data yang digunakan untuk melakukan penelitian ini diperoleh dari komentar (ulasan) pelanggan Hanatex yang disampaikan melalui chat Whatsapp. Pengambilan data mentah atau dokumen teks (ulasan) diambil sebanyak 125 ulasan. Ulasan yang ada kemudian di filter secara manual. Data yang diambil adalah data mengandung keluhan mengenai produk. Sedangkan sisanya sebanyak 25 ulasan tidak digunakan karena ulasanya tidak termasuk kedalam ulasan yang mengandung keluhan mengenai produk.

Tabel 1 Data Komentar Pelanggan.

Pelanggan	Komentar
0888690xxxx	Mas kaji benangnya ukurannya kebesaran
0815754xxxx	Bos Benangnya yang satu bal mudah putus tak retur ya
0815781xxxx	Kiriman benang minggu kemarin ada yang kecampuran dengan benang TC 3 cones
0888666xxxx	Benang 30 Ct kecampuran sama 40 TC
<i>dst...</i>	

### 3.2 Text Processing

Setelah dokumen teks diperoleh, selanjutnya data tersebut masuk ketahap Text Preprocessing. Text Preprocessing bertujuan untuk mengolah dataset agar dapat dibaca oleh algoritma dan mengekstraksi informasi yang di inginkan. Text preprocessing akan menghasilkan data bersih atau clean-dataset yang berisi informasi spesifik. Proses ini meliputi:

#### 1. *Normalize Case*

Pada tahap *Normalize Case* akan melakukan perubahan teks pada Dokumen Teks yang digunakan, untuk perubahan dokumen teks yang dilakukan adalah melakukan konversi teks yang terdapat huruf besar di dokumen teks akan dirubah menjadi huruf kecil semua.

Tabel 2 Contoh Tahap Normalize Case

Dokument Teks (Input)	Normalize Case (Output)
Kiriman benang minggu kemarin ada yang kecampuran dengan benang tc 3 kones	kiriman benang minggu kemarin ada yang kecampuran dengan benang tc 3 kones
Benang 30 Ct kecampuran sama 40 TC	benang 30 ct kecampuran sama 40 tc
benang 40 Ct kecampuran klenteng ada 5 kones	benang 40 ct kecampuran klenteng ada 5 kones
<i>dst...</i>	

#### 2. *Tokenizing*

Sedangkan pada tahap *Tokenizing* adalah tahapan untuk melakukan pemotongan string input berdasarkan tiap kata yang menyusunnya, atau melakukan perubahan dari kalimat yang ada menjadi kata-kata tersendiri.

Tabel 3 Contoh Tahap *Tokenizing*

Dokument Teks (Input)	Tokenizing (Output)
kiriman benang minggu kemarin ada yang kecampuran dengan benang tc 3 kones	kiriman benang minggu kemarin ada yang kecampuran dengan benang tc 3 kones
Benang 30 Ct kecampuran sama 40 TC	benang 30 ct kecampuran sama 40 tc
<i>dst...</i>	

#### 3. *Filtering*

Untuk tahap *Filtering* adalah tahap mengambil kata-kata penting dari hasil dokument teks (*token*) yang sudah ada. Tahap *Filtering* ada dua (2) yaitu: *filter tokens by lenght*, algoritma ini digunakan untuk membuang kata-kata yang ukurannya terlalu sedikit atau terlalu banyak, dalam penelitian ini kata-kata yang akan dibuang adalah kata kata yang memiliki ukuran kurang dari 4 dan lebih dari 25 kata. Sedangkan *algoritma filter stopwords* digunakan untuk membuang kata-kata tidak penting yang ada di dokument teks.

Tabel 4 Contoh filter token by length

yg	bli	bos	ct
dah	dgn	di	ga
e	pak	kok	mas
nya	oke	om	pas
sm	sy	tdk	tp
dst..	dst...	dst..	dst..

Tabel 4 merupakan kata-kata yang akan dibuang, karena ukuran kata terlalu kecil yaitu kurang dari 4, untuk kata-kata yang

lebih dari 25 pada dokumen teks yang digunakan dalam penelitian ini tidak ada.

Tabel 5 Contoh Tahap Filtering (Stop Words List).

dari	benar	cukup	enggak	jelas
atau	berada	cuma	entah	jika
adalah	berapa	dibuat	guna	jumlah
adanya	beri	dijawab	hampir	kalau
adapun	berupa	agar	harus	kamu
agak	boleh	disebut	inginkan	keluar
akhir	bukan	diucapkan	itulah	khususnya
...	...	...	...	...

Tabel 5 merupakan kata-kata yang akan dibuang karena kata-kata tersebut tidak memiliki arti yang penting dalam dokumen text yang menyangkut keluhan.

4. *Stemming*

Tahap *Stemming* adalah melakukan proses penghapusan kata-kata yang mempunyai awalan ataupun akhiran disetiap kata-katanya, proses ini juga disebut juga sebagai proses untuk memperoleh kata dasar dari kata yang digunakan disetiap kata-kata yang ada, atau mencari *root* kata dari tiap kata hasil filtering.

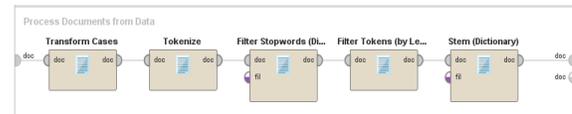
Tabel 6 Contoh Tahap *Stemming*

Teks Sebelum <i>Stemming</i> (Input)	Teks Sebelum <i>Stemming</i> (Output)
kualitasnya	kualitas
ukurannya	ukuran
Campuran	campur
Campurannya	campur
kecampuran	campur
...	...

Berdasarkan hasil *Text Preprocessing* data kemudian akan dilanjutkan ke tahap proposed method yang merupakan proses pembentukan cluster menggunakan algoritma k-means untuk

mendapatkan segmen-segmen keluhan pelanggan.

Proses Text Preprocessing yang telah dilakukan pada experiment ini dapat diketahui pemodelannya pada gambar gambar 3.1



Gambar 3 Proses Text Processing.

3.3 Proposed Method

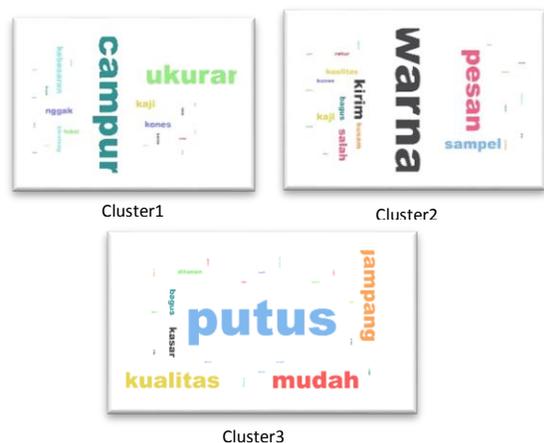
Tahap ini merupakan tahap pembentukan model untuk menentukan segmen-segmen keluhan dari pelanggan yang telah disampaikan melalui chat di Whatsapp mengenai suatu produk yang dijual oleh Hanatex. Dalam proses ini akan dilakukan klasterisasi keluhan pelanggan menggunakan algoritma k-means, dalam penentuan nilai cluster digunakan elbow methode untuk menentukan nilai k yang optimal. Nilai k optimal yang ditentukan berdasarkan elbow methode yaitu 3.

Hasil eksperimen yang terbentuk dari model Analisa keluhan pelanggan dengan metode k-means terbentuk kedalam 3 segmen yaitu seperti pada gambar 5.2

Attribute	cluster_0 ↓	cluster_1	cluster_2
campur	0.233	0.008	0
ukuran	0.203	0	0.022
nggak	0.080	0	0
lones	0.074	0.033	0.039
benang	0.069	0.043	0.066
kaji	0.061	0.086	0.020
kebesaran	0.058	0	0
menteng	0.049	0	0
tebal	0.049	0	0
kecewa	0.039	0	0
bagus	0.037	0.063	0.090
opsional	0.036	0	0

Gambar 4 Cluster yang terbentuk dari algoritma k-means

Setiap cluster yang terbentuk akan divisualisasikan berdasarkan frekuensi kata-kata ke dalam *Wordcloud* untuk mengetahui keluhan keluhan pelanggan tersebut ada disegmen apa. Bentuk dari visualisasi cluster yang terbentuk ditunjukkan pada gambar 3.3.



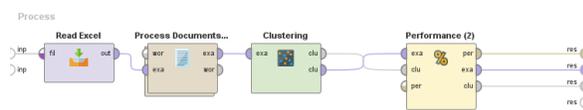
Gambar 5 Wordcloud cluster yang terbentuk

### 3.4 Evaluasi (hasil analisa)

Tahap evaluasi merupakan tahap pengujian dari performance algoritma dari suatu model yang terbentuk dari eksperimen mengenai segmentasi keluhan pelanggan dengan algoritma k-means. Proses evaluasi menggunakan bantuan software Rapid Miner 9.10.

Dalam proses evaluasi digunakan DBI (Davies Bouldin Index) yang berfungsi untuk mengetahui performance dari algoritma k-means dengan metode pendekatan berdasarkan penentuan nilai kluster yang optimal dengan melihat jarak antara kluster yang satu dengan yang lainnya dan pada waktu yang sama. Jumlah kluster yang terbaik yaitu yang memiliki nilai DBI minimum.

Berikut proses evaluasi dari penerapan algoritma k-means dalam menentukan segmentasi keluhan pelanggan yang dilakukan menggunakan bantuan software Rapid Miner 9.10, seperti terlihat pada gambar 5.3.



Gambar 6 Evaluasi Algoritma k-means dengan DBI

Hasil evaluasi dari algoritma k-means dalam mengcluster keluhan pelanggan dengan DBI diperoleh nilai sebesar 0.112 pada jumlah kluster 3. Artinya bahwa kemiripan data dokumen text pada anggota dalam cluster sebesar 88,8% mirip.

## 4. Kesimpulan dan Saran

Berdasarkan hasil clusterisasi keluhan pelanggan dengan metode k-means dengan menggunakan 100 Komentar (ulasan) dari pelanggan, dapat disimpulkan:

1. Segmentasi yang terbentuk dari clusterisasi keluhan pelanggan dengan metode k-means ada 3 segmentasi yaitu segmentasi pertama keluhan mengenai ukuran benang yang tercampur, segmentasi yang kedua mengenai warna benang yang dipesan tidak sesuai sample dan segmentasi yang ke tiga mengenai kualitas benang mudah putus
2. Keluhan pelanggan yang paling banyak yaitu pada segmen kedua yaitu mengenai warna benang yang dipesan tidak sesuai dengan sample mencapai 45%, kemudian disusul keluhan mengenai ukuran benang tercampur dengan prosentase 31% pada segmen satu, sedangkan keluhan yang paling sedikit yaitu di segmen ketiga yaitu mengenai kualitas benang yang mudah putus sebanyak 24%.

Berdasarkan hasil evaluasi menggunakan DBI yaitu sebesar 0.112, artinya performance dari algoritma k-means dengan menggunakan nilai cluster 3 dalam melakukan segmentasi dapat disimpulkan baik karena 88,8% kemiripan data dalam kelompok mirip.

Meskipun dalam penelitian ini hasil segmentasi yang didapatkan adalah sesuai, akan tetapi untuk penelitian berikutnya disarankan perlu ditambahkan metode yang berfungsi untuk cleaning data pada text processing untuk menghilangkan kata kata singkatan maupun kata-kata yang tidak baku agar dapat meningkatkan performance dan akurasi metode.

## DAFTAR PUSTAKA

- Al'asqolaini, M Z. 2019. "Penanganan Keluhan Pelanggan Dan Kualitas Layanan Pengaruhnya Terhadap Loyalitas Pelanggan." *Jurnal Ekonomi Dan Kewirausahaan* ....  
<http://journal.stienugresik.ac.id/index.php/jek/article/view/77%0Ahttp://journal.stienugresik.ac.id/index.php/jek/article/download>

- /77/48.
- Amalia, Nurul, Nur Ika Royanti, and Bambang Ismanto. 2023. "Analisa Sentimen Pelanggan Pada Review Belanja Online Berbasis Text Mining Menggunakan Metode K-Means" 4 (4): 1441–47. <https://doi.org/10.47065/josh.v4i4.3781>.
- Ani, Isro, and Wahyu Jatmiko. 2023. "Pengaruh Kualitas Pelayanan Dan Penanganan Komplain Terhadap Loyalitas Pelanggan PT. First Media Di Sidoarjo." *Jurnal Ilmiah Administrasi Bisnis Dan Inovasi* 6 (2): 165–82. <https://doi.org/10.25139/jiabi.v6i2.5325>.
- Ariyanti, Dyah, and Kurnia Iswardani. 2020. "Teks Mining Untuk Klasifikasi Keluhan Masyarakat Pada Pemkot Probolinggo Menggunakan Algoritma Naïve Bayes." *Jurnal IKRA-ITH Informatika* 4 (3): 125–32.
- Ashma Nurmeila, Sheila, Wina Witanti, and Puspita Sabrina Nurul. 2020. "Segmentasi Pelanggan Berdasarkan Keluhan Dengan Menggunakan K-Means Cluster Analysis Pada PT Infomedia Nusantara." *Prosiding Seminar Nasional Sistem Informasi Dan Teknologi (SISFOTEK)*, 276–80.
- Kusumo, Siheruadi, and Diah Riandadari. 2017. "Analisis Kualitas Produksi Benang Dengan Metode Six Sigma Pada Pt. Sejati Jaya." *Jurnal Pendidikan Teknik Mesin UNESA* 6 (01): 28–34.
- Napitupulu, Ferdinand. 2019. "Pengaruh Harga Dan Kualitas Produk Terhadap Kepuasan Pelanggan Pada PT. Ramayana Lestari Sentosa." *Kinerja* 16 (1): 1–9.
- Putro, Shandy Widjoyo. 2014. "Pengaruh Kualitas Layanan Dan Kualitas Produk Terhadap Kepuasan Pelanggan Dan Loyalitas Konsumen Restoran Happy Garden Surabaya." *Jurnal Manajemen Pemasaran* 2 (1): 1–9. <http://publication.petra.ac.id/index.php/management-pemasaran/article/view/1404>.
- Riyanto, Eko. 2020. "Analisis Profitabilitas Pada Industri Tekstil Dan Produk Tekstil Di Indonesia." *Jurnal Dinamika Ekonomi Pembangunan* 3 (2): 77–89. <https://doi.org/10.14710/jdep.3.2.77-89>.
- Saputra, Try Iryanto, and Rini Arianty. 2019. "Implementasi Algoritma K-Means Clustering Pada Analisis Sentimen Keluhan Pengguna Indosat." *Jurnal Ilmiah Informatika Komputer* 24 (3): 191–98. <https://doi.org/10.35760/ik.2019.v24i3.2361>.