# **LAPORAN PROYEK AKHIR**

**PRAKTIKUM DATA SCIENCE**

**ANALISIS SENTIMENT TERHADAP**

**PAKAIAN WANITA MENGGUNAKAN METHODE LDA ( *Latent Dirichlet Allocation* )**



|  |  |
| --- | --- |
| Fatin Luthfi Salman | 123200071 |
| Shakira Luthfiani A | 123200165 |

**PROGRAM STUDI INFORMATIKA**

**JURUSAN INFORMATIKA**

**FAKULTAS TEKNIK INDUSTRI**

**UNIVERSITAS PEMBANGUNAN NASIONAL “VETERAN” YOGYAKARTA**

**2022**

# **PENDAHULUAN**

Pakaian adalah kebutuhan pokok manusia selain makanan dan tempat tinggal (rumah). Manusia membutuhkan pakaian untuk melindungi dan menutup dirinya. Namun seiring dengan perkembangan kehidupan manusia, pakaian juga digunakan sebagai simbol status, jabatan, ataupun kedudukan seseorang yang memakainya. Perkembangan mode dan jenis-jenis pakaian tergantung pada adat-istiadat, kebiasaan, dan budaya, pada suatu budaya juga terdapat ciri khas yang berbeda dengan budaya lain. Dengan hadirnya media sosial saat ini membuat penyebaran perkembangan mode pakaian menjadi sangat cepat.

Berbagai toko online mulai memanfaatkan momentum ini untuk menjaring banyak pelanggan baru, salah satunya pakaian wanita. Di masa pandemi, belanja online menjadi pilihan utama banyak orang. 92% mencoba cara baru untuk berbelanja, 57% berbelanja secara digital, dan 48% menggunakan aplikasi pengambilan dan pengiriman bahan makanan (Ferry Kusnowo, 2020). Banyak pelanggan baru yang tidak mengetahui kualitas produk yang dijual di toko online, sehingga membuat pelanggan baru mengecek *review* produk sebelum membeli. Nah, untuk memudahkan pelanggan baru menarik kesimpulan dari *review* yang mereka lihat, maka dilakukan pengujian data mining berupa *review* atau *review* pakaian wanita.

* 1. **Analisis Sentiment**

Analisis sentiment atau biasa dikenal dengan *opinion mining* adalah studi komputasi untuk mengidentifikasi dan mengungkapkan opini, sentiment, evaluasi, sikap, sentiment, subjektivitas, penilaian, atau pendapat yang terkandung dalam teks. Analisis sentiment juga merupakan metode untuk mengekstraksi data opini, memahami dan mengolah data teks secara otomatis untuk melihat sentiment yang terkandung dalam opini.

Ulasan adalah bagian dari kecerdasan manusia dan memainkan peran penting dalam pengambilan keputusan manusia. Manusia sebagian besar mengabaikan emosi dalam interaksi komputer karena perannya yang bias dalam komunikasi manusia ke manusia sehari-hari.

* 1. ***Preprocessing***

*Preprocessing* merupakan salah satu tahapan penting dalam proses data mining. Data yang digunakan dalam proses penambangan tidak selalu dalam kondisi pemrosesan yang ideal. Terkadang terdapat berbagai permasalahan pada data yang mengganggu hasil dari proses mining itu sendiri, seperti *missing value*, *redundant data, outlier,* atau *format* data yang tidak sesuai dengan sistem. Oleh karena itu, untuk mengatasi permasalahan tersebut diperlukan tahapan preprocessing.

Preprocessing adalah salah satu tahapan yang menghilangkan masalah yang dapat mengganggu hasil pengolahan data. Dalam kasus klasifikasi dokumen menggunakan data tipe teks, beberapa proses biasanya dilakukan, termasuk pelipatan kasus, penyaringan (penghapusan tanda baca), penghapusan kata berhenti, *stemming, tokenisasi*, dll.

* 1. ***Shiny***

*Shiny* adalah paket dalam R yang memungkinkan pengguna membangun aplikasi web interaktif. *Shiny* menggabungkan kekuatan komputasi statistik R dengan interaksinya dengan web interaktif modern. Struktur Shiny terdiri dari 4 blok atau komponen yaitu *global, ui, server* dan *run-app*.

Blok global berisi paket (perpustakaan) yang diperlukan agar jaringan berfungsi dengan baik. UI adalah fitur yang menentukan tampilan dan nuansa web dari aplikasi yang akan dijalankan. Fungsinya berisi semua input dan output yang akan ditampilkan di aplikasi. Server adalah fungsi yang mendefinisikan logika kerja analisis dari sisi server aplikasi. Run-App adalah fungsi aplikasi untuk memanggil UI dan Server untuk menjalankan aplikasi.

* 1. **Algoritma *Latent Dirichlet Allocation* (LDA)**

Analisis sentiment dapat membantu dalam mengekstrasi pendapat yang berasal dari sebuah dokumen, komentar, *review* produk, dan data-data lainnya sehingga akan lebih mudah dalam melakukan pemantauan. Salah satu metode untuk melakukan analisis sentiment adalah *Latent Dirichlet Allocation* (LDA) yang mampu digunakan untuk mengekstraksi topik dari kumpulan dokumen komentar, dimana topik tersebut direpresentasikan sebagai kemunculan kata-kata dengan probabilitas topik yang berbeda. Oleh karena itu diperlukan representasi data dalam bentuk visual yang mudah dipahami daripada teks dan tabel. Salah satu bentuk visualisasi data adalah *wordcloud* yang memberikan gambaran visual frekuensi kemunculan kata.

* 1. **WordCloud**

*Wordcloud* (juga dikenal sebagai awan teks atau awan tag) adalah cara untuk memvisualisasikan data teks. Grafik ini populer dalam penambangan teks karena mudah dipahami. Dengan menggunakan *wordcloud*, gambaran frekuensi kata dapat ditampilkan dalam bentuk yang menarik namun tetap informatif. Semakin sering sebuah kata digunakan, semakin besar kata tersebut akan muncul di wordcloud.

Pada projek ini, kami menggunakan label rating sebagai representasi dari klasifikasi jenis komentar. Berikut ini prosesnya :

|  |  |  |  |  |  |
| --- | --- | --- | --- | --- | --- |
| No | Id | Review Text | Rating | Positive Feedback Count | Class Name |
| 1 | 1077 | I had such high hopes for this dress and really wanted it to work for me. i initially ordered the petite small (my usual size) but i found this to be outrageously small. so small in fact that i could not zip it up! i reordered it in petite medium, which was just ok. overall, the top half was comfortable and fit nicely, but the bottom half had a very tight under layer and several somewhat cheap (net) over layers. imo, a major design flaw was the net over layer sewn directly into the zipper - it | 3 | 0 | Dresses |
| 2 | 1049 | I love, love, love this jumpsuit. it's fun, flirty, and fabulous! every time i wear it, i get nothing but great compliments! | 5 | 0 | Pants |
| 3 | 1077 | I love this dress. i usually get an xs but it runs a little snug in bust so i ordered up a size. very flattering and feminine with the usual retailer flair for style. | 5 | 0 | Dresses |
| 4 | 767 | Bought the black xs to go under the larkspur midi dress because they didn't bother lining the skirt portion (grrrrrrrrrrr). my stats are 34a-28/29-36 and the xs fit very smoothly around the chest and was flowy around my lower half, so i would say it's running big. the straps are very pretty and it could easily be nightwear too. i'm 5'6" and it came to just below my knees. | 5 | 0 | Intimates |
| 5 | 949 | I have been waiting for this sweater coat to ship for weeks and i was so excited for it to arrive. this coat is not true to size and made me look short and squat. the sleeves are very wide (although long). as a light weight fall coat the sleeves don't need to be as wide because you wouldn't be layerng too much underneath. the buttons need to be moved at least three inches in for a nicer fit. i thought about redoing the buttons myself but the sleeves looked even more out of proportion with a tigh | 2 | 0 | Sweaters |

Setelah data diperoleh data kemudian diklasifikasi sebagai komentar positif atau negatif berdasarkan ratingnya. Untuk rating kurang dari atau sama dengan 3 maka akan diklasifikasikan sebagai komentar negatif, sedangkan untuk rating lebih dari 3 akan diklasifikasikan sebagai komentar positif.

|  |  |  |  |  |  |  |
| --- | --- | --- | --- | --- | --- | --- |
| No | Id | Review Text | Rating | Positive Feedback Count | Class Name | Klasifikasi |
| 1 | 1077 | I had such high hopes for this dress and really wanted it to work for me. i initially ordered the petite small (my usual size) but i found this to be outrageously small. so small in fact that i could not zip it up! i reordered it in petite medium, which was just ok. overall, the top half was comfortable and fit nicely, but the bottom half had a very tight under layer and several somewhat cheap (net) over layers. imo, a major design flaw was the net over layer sewn directly into the zipper - it | 3 | 0 | Dresses | Negatif |
| 2 | 1049 | I love, love, love this jumpsuit. it's fun, flirty, and fabulous! every time i wear it, i get nothing but great compliments! | 5 | 0 | Pants | Positif |
| 3 | 1077 | I love this dress. i usually get an xs but it runs a little snug in bust so i ordered up a size. very flattering and feminine with the usual retailer flair for style. | 5 | 0 | Dresses | Positif |
| 4 | 767 | Bought the black xs to go under the larkspur midi dress because they didn't bother lining the skirt portion (grrrrrrrrrrr). my stats are 34a-28/29-36 and the xs fit very smoothly around the chest and was flowy around my lower half, so i would say it's running big. the straps are very pretty and it could easily be nightwear too. i'm 5'6" and it came to just below my knees. | 5 | 0 | Intimates | Positif |
| 5 | 949 | I have been waiting for this sweater coat to ship for weeks and i was so excited for it to arrive. this coat is not true to size and made me look short and squat. the sleeves are very wide (although long). as a light weight fall coat the sleeves don't need to be as wide because you wouldn't be layerng too much underneath. the buttons need to be moved at least three inches in for a nicer fit. i thought about redoing the buttons myself but the sleeves looked even more out of proportion with a tigh | 2 | 0 | Sweaters | Negatif |

# **METODE**

Proses penelitian menggunakan teknik unsupervised learning yang dijelaskan oleh (*Vishwanathan* dan *Smola*, 2010). unsupervised learning merupakan salah satu tipe algoritma *machine learning* yang digunakan untuk menarik kesimpulan dari dataset. Metode ini hanya akan mempelajari suatu data berdasarkan kedekatannya saja atau yang biasa disebut dengan *clustering*. Proses *clustering* menggunakan metode *unsupervised learning* untuk menghasilkan klaster positif dan negatif dari *review* pakaian wanita.

**2.1 Pengumpulan Data**

Pengumpulan data *training review women’s clothing e-commerce* menggunakan parameter yang sudah tersedia dari website https://www.kaggle.com untuk mendapatkan *review* kategori positif atau negatif. Data yang telah diambil disimpan dalam file data womensclothing.csv dan akan dilanjutkan proses preprocessing data.

**2.2 Preprocessing**

*Preprocessing* dilakukan mulai dari proses *cleaning* untuk mempermudah proses analisis seperti menghilangkan tanda baca atau *punctuation* dan juga *hashtag* atau *mention.* Kemudian proses merubah semua huruf menjadi huruf kecil *tolower* dan terakhir yaitu proses menghapus *stopword* atau kata yang tidak sesuai dengn kaidah Bahasa Indonesia, mengubah kata yang terdapat imbuhan menjadi kata dasar *stemming*, mengubah kata yang mempunyai makna menjadi hanya sebuah kata *lemmatisasi* dan terakhir mengubah kata yang *typo* menjadi kata baku *slankword* sehingga setelah *preprocessing* selesai maka data sudah dapat diolah.

**2.3 Topic *Modeling Latent Dirichlet Allocation* (LDA)**

*Latent Dirichlet Allocation* (LDA) digunakan untuk menentukan topik secara otomatis dari sekumpulan ulasan *review*. Ulasan *review* yang diteliti memiliki struktur tersembunyi (hidden structure) berupa topik, distribusi topik per ulasan review, dan penentuan topik per kata dalam setiap ulasan *review*. LDA menggunakan kumpulan ulasan *review* tersebut untuk inference struktur topik tersembunyi. Jumlah topik yang akan dihasilkan telah ditentukan sebelum proses LDA dilaksanakan.

Setelah model topik LDA dibuat, sebuah ulasan *review* dapat ditentukan distribusi topiknya yang mendeskripsikan kumpulan kata dalam ulasan *review*. LDA menggunakan asumsi *bag of words*, yaitu urutan kemunculan kata dalam dokumen diabaikan.

**2.4WordCloud**

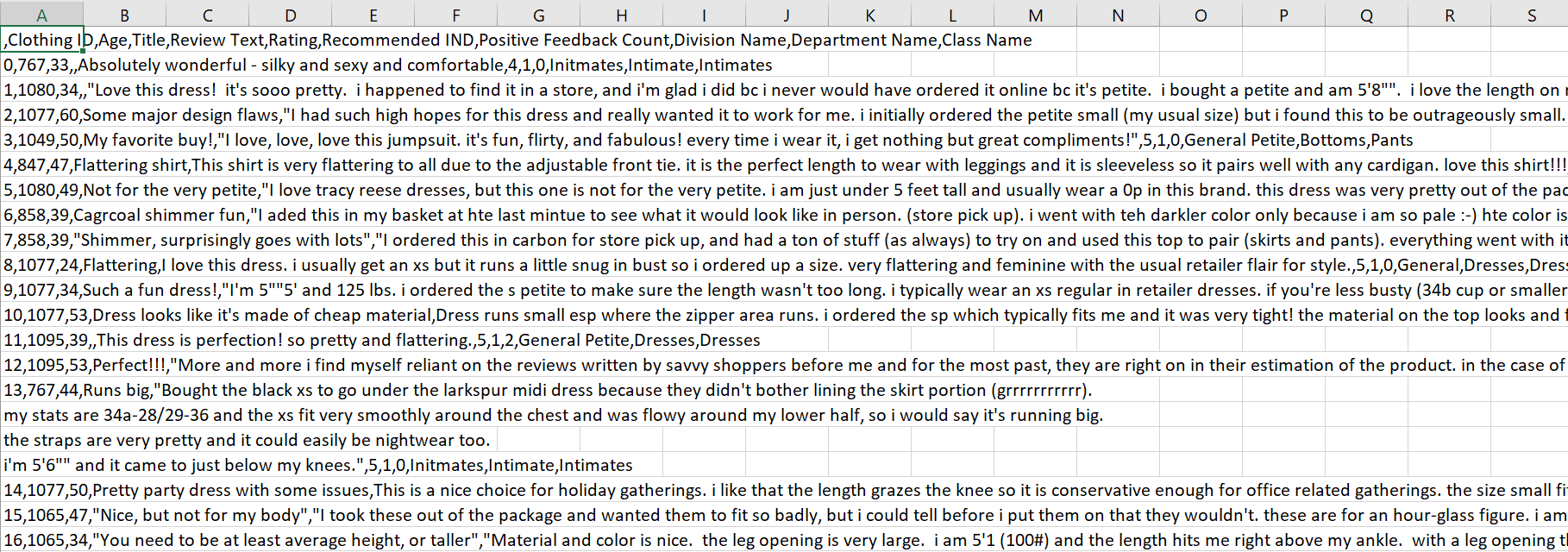
Pembuatan *wordcloud* menggunakan package *wordcloud* dan dapat diinstal dalam R-studio dari repositori *CRAN* kemudian dikerjakan dalam dua tahap, yaitu persiapan data dan pembuatan *wordcloud*. Data yang diambil berupa kumpulan kata negatif dan positif yang sudah diklasifikasikan. Kemudian tampilan *wordcloud* kata negatif berwarna hitam dan berada di bagian atas sedangkan kata positif berwarna biru yang berada dibagian bawah. Total kata yang ditampilkan dalam word cloud berisi 200 kata negatif dan positif.

# **HASIL DAN PEMBAHASAN**

Pada Penelitian proyek ini data diolah dengan algoritma *Latent Dirichlet Allocation* (LDA). Berikut adalah hasil dan pembahasan dari langkah uji coba yang dilakukan pada analisis :

**3.1 Pengumpulan data**

Data didapatkan dari kaggle, Data yang diambil berformat csv yang dilakukan pengumpulan data menggunakan library (vroom) dengan aplikasi RStudio, didapatkan 23.486 komentar.



**Gambar 1.** Hasil Crawling Data

**3.2 Preprocessing**

Setelah mendapatkan data kotor dari *Kaggle*, langkah selanjutnya adalah preprocessing. Melalui tahapan ini, dataset kotor diproses melalui berbagai proses seperti: penghapusan URL, "\n", koma, RT, titik dua, titik koma, titik dua, dan simbol. Tujuan dari proses ini sendiri adalah untuk menghasilkan kumpulan data yang bersih. Berikut adalah contoh hasil dari 5 data yang telah dilakukan tahap *preprocessing.*

**Tabel 1.** Hasil Preprocessing

|  |  |
| --- | --- |
| NO | Ulasan |
| 1 | Absolutely wonderful - silky and sexy and comfortable |
| 2 | Love this dress! it's sooo pretty. i happened to find it in a store, and i'm glad i did bc i never would have ordered it online bc it's petite. i bought a petite and am 5'8". i love the length on me- hits just a little below the knee. would definitely be a true midi on someone who is truly petite. |
| 3 | I had such high hopes for this dress and really wanted it to work for me. i initially ordered the petite small (my usual size) but i found this to be outrageously small. so small in fact that i could not zip it up! i reordered it in petite medium, which was just ok. overall, the top half was comfortable and fit nicely, but the bottom half had a very tight under layer and several somewhat cheap (net) over layers. imo, a major design flaw was the net over layer sewn directly into the zipper - it c |
| 4 | I love, love, love this jumpsuit. it's fun, flirty, and fabulous! every time i wear it, i get nothing but great compliments! |
| 5 | Dress runs small esp where the zipper area runs. i ordered the sp which typically fits me and it was very tight! the material on the top looks and feels very cheap that even just pulling on it will cause it to rip the fabric. pretty disappointed as it was going to be my christmas dress this year! needless to say it will be going back. |

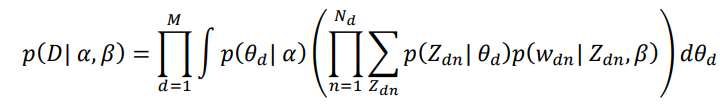
Berikut adalah contoh hasil dari data yang sudah dilakukan pelabelan menggunakan algoritma pencocokan sederhana, dapat dilihat pada Tabel 2.

**Tabel 2.** Hasil Scoring

|  |  |  |  |
| --- | --- | --- | --- |
| NO | Ulasan | Age | Klasifikasi |
| 1 | Absolutely wonderful - silky and sexy and comfortable | 33.00 | Positif |
| 2 | Love this dress! it's sooo pretty. i happened to find it in a store, and i'm glad i did bc i never would have ordered it online bc it's petite. i bought a petite and am 5'8". i love the length on me- hits just a little below the knee. would definitely be a true midi on someone who is truly petite. | 34.00 | Positif |
| 3 | I had such high hopes for this dress and really wanted it to work for me. i initially ordered the petite small (my usual size) but i found this to be outrageously small. so small in fact that i could not zip it up! i reordered it in petite medium, which was just ok. overall, the top half was comfortable and fit nicely, but the bottom half had a very tight under layer and several somewhat cheap (net) over layers. imo, a major design flaw was the net over layer sewn directly into the zipper - it c | 60.00 | Negatif |
| 4 | I love, love, love this jumpsuit. it's fun, flirty, and fabulous! every time i wear it, i get nothing but great compliments! | 50.00 | Positif |
| 5 | Dress runs small esp where the zipper area runs. i ordered the sp which typically fits me and it was very tight! the material on the top looks and feels very cheap that even just pulling on it will cause it to rip the fabric. pretty disappointed as it was going to be my christmas dress this year! needless to say it will be going back. | 53.00 | Negatif |

**3.3 Menghitung Probabilitas Kata**

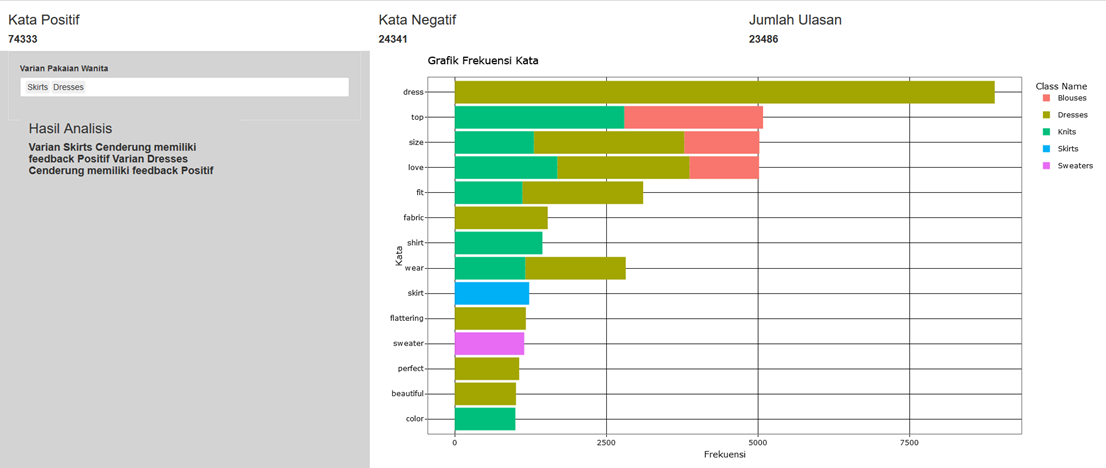
Berikut merupakan perhitungan probabilitas dari sebuah *corpus* berdasarkan notasi yang telah dijelaskan



Dapat dilihat bahwa pada notasi **β** mendeskripsikan topik, dimana pada setiap **β** merupakan distribusi dari sejumlah kata. Pada Variabel **θd** adalah variabel level dokumen dengan satu kali sampel per dokumen yang merepresentasikan proporsi topik untuk dokumen ke d. Pada notasi zdn dan wdn merupakan representasi variabel di level kata dengan satu kali sampel untuk masing-masing kata pada setiap dokumen.

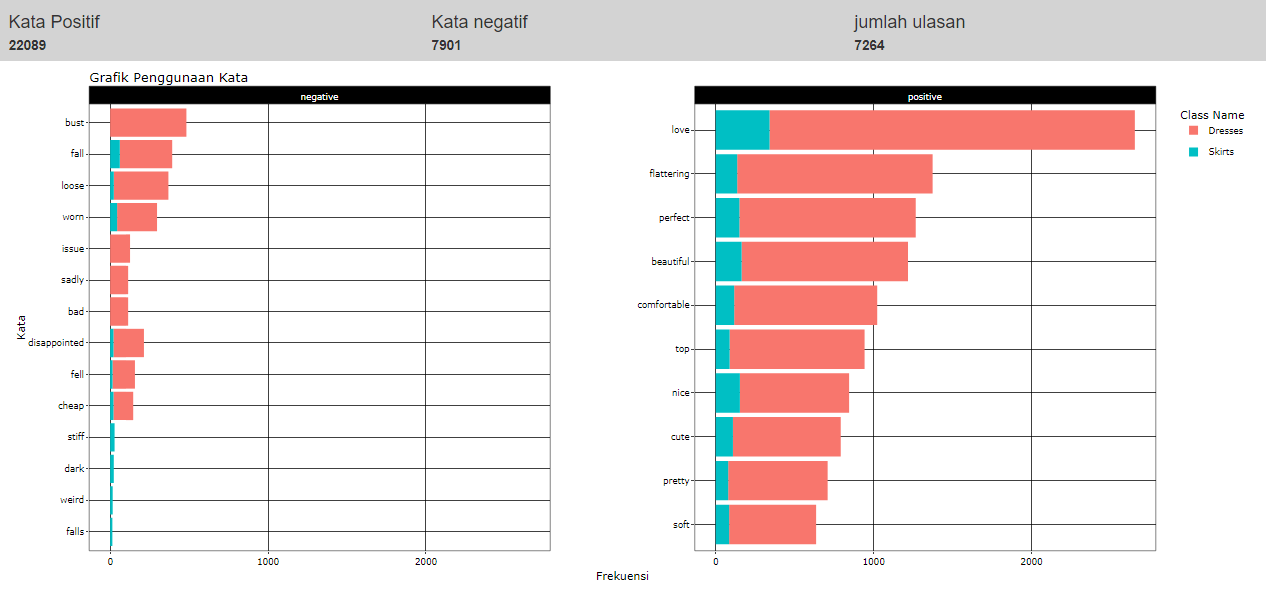
**3.4 Tampilanshiny**

Setelah Seluruh data sudah dibersihkan dan difilter, kita akan mendapatkan grafik *interface* hasil Sentimentt Analysis pakaian wanita yang sudah dikategorikan menjadi beberapa variasi seperti (*blouses, dresses, knits, skirts, sweaters*) dan kesimpulan hasil analisis sesuai dengan varian yang dipilih, yang dapat dilihat pada gambar dibawah :



**Gambar 2.** *Word Frequency Graphic*

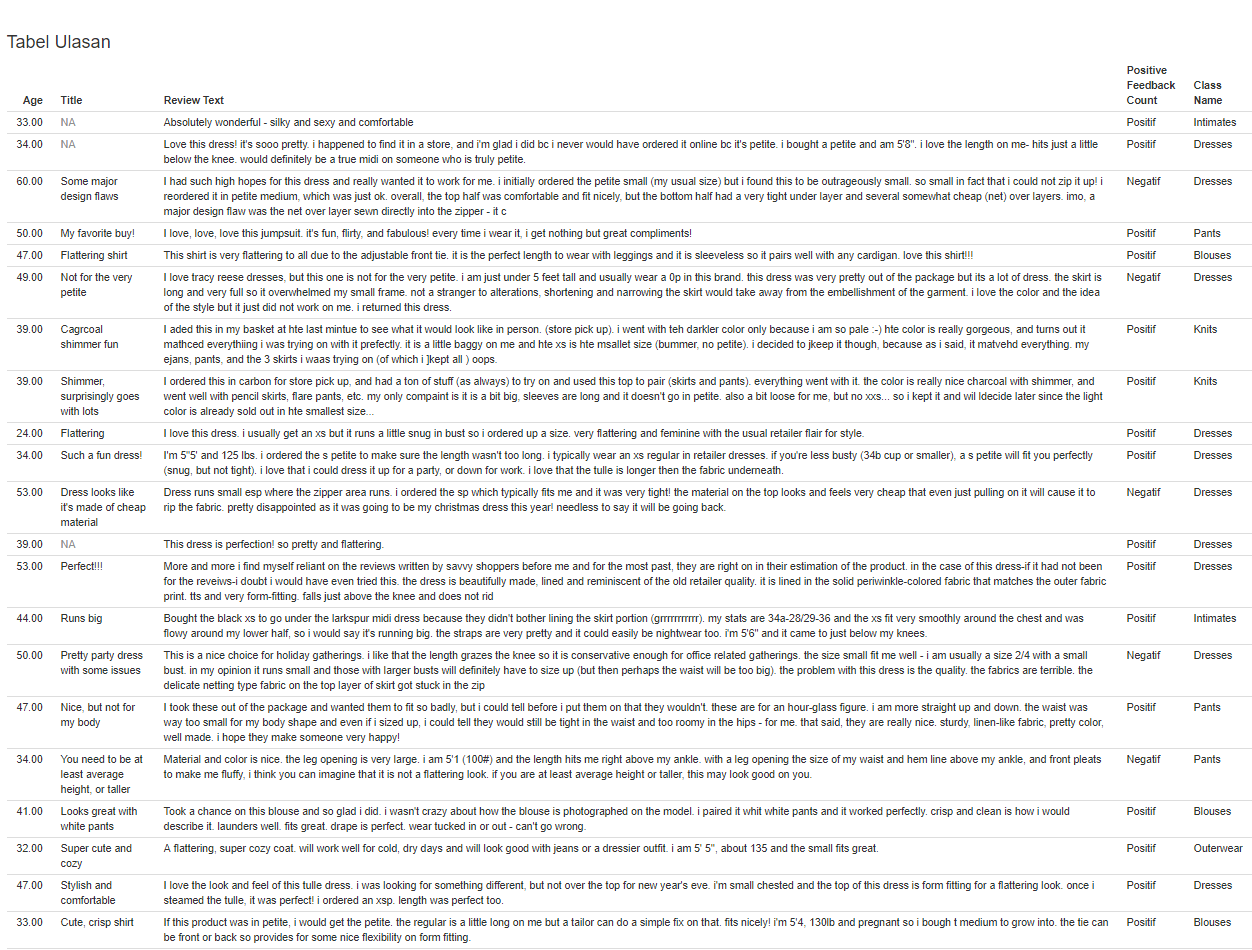
Hasil dari polaritas yang diambil berdasarkan ulasan review dikategorikan menjadi dua bentuk analisis polaritas seperti: Positif dan Negatif. Sehingga dapat Menunjukkan jumlah masing-masing kategori polaritas. Dari hasil yang didapatkan sebelumnya disajikan ke dalam bentuk grafik interface yang terlihat pada Gambar 3. untuk mempermudah proses pembacaan hasil akhir.



**Gambar 3**. Word Usage Graohic

**3.4 Tampilan Shiny Tabel Data**

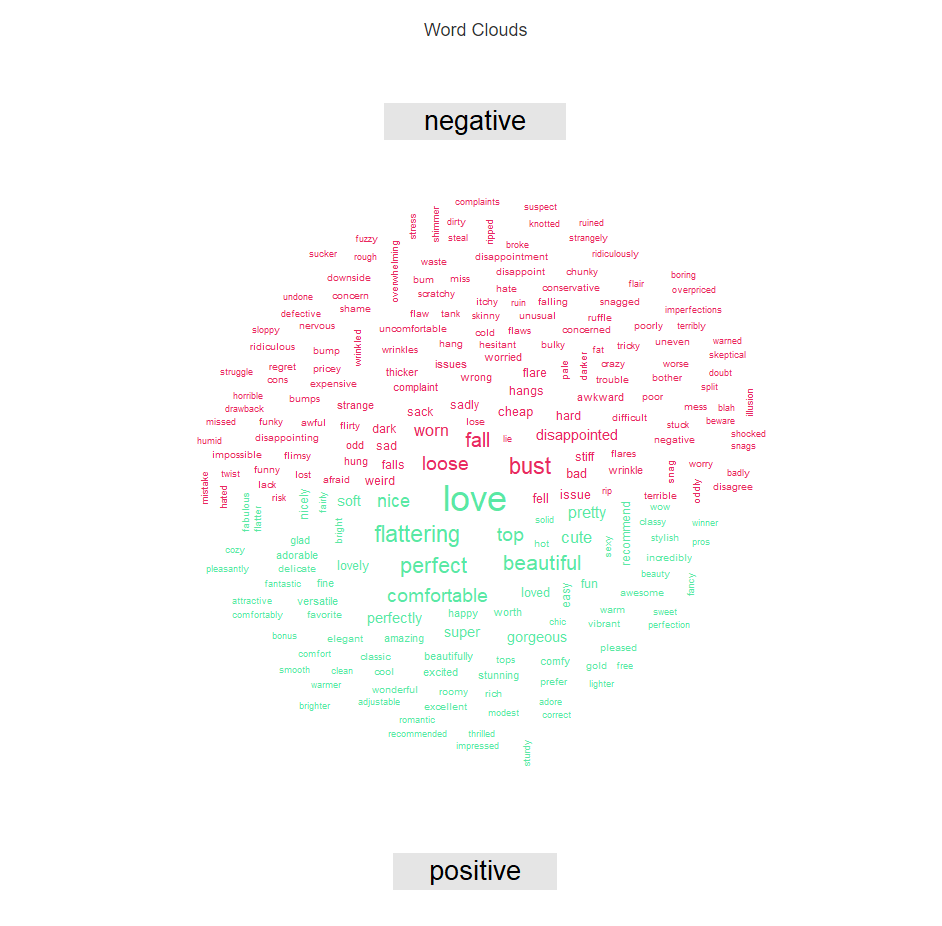
*Shiny* juga dapat menampilkan tabel dari data yang sudah diolah. Data yang sudah melalui tahap *preprocessing* pada tampilan dalam bentuk tabel dan dimuat dalam beberapa halaman karena terdapat banyak data.



**Gambar 4.** Tabel Data

**3.5 Tampilan Shiny Wordcloud**

*Shiny* juga menampilkan *wordcloud* dari hasil analisis. Hasil dari *wordcloud* diambil berdasarkan ulasan review yang dikategorikan menjadi dua bentuk analisis yaitu kata negatif dan positif seperti gambar 5. Semakin besar kata dalam *wordcould* yang ditampilkan maka kata tersebut adalah kata yang paling banyak atau sering digunakan dalam ulasan review.



**Gambar 5.** Wordcloud

**3.6 Listing Program**

|  |
| --- |
| ---  title: "Project Akhir Prak DS"  author: "Fatin Luthfi Salman 123200071 || Shakira Luthfiani A 123200165"  date: "11/20/2022"  output: html\_document  ---  ```{r setup, include=FALSE}  knitr::opts\_chunk$set(echo = TRUE)  ```  \*\*global\*\*  ```{r global}  # Daftar library Yang digunakan  library(dplyr)#untuk memanipulasi data  library(shiny)#untuk membuat web sederhana  library(shinyWidgets)  library(tidyverse)#untuk mengolah data  library(vroom)#untuk import data dari file  library(here)#untuk menemukan file  library(tidytext)#untuk text mining  library(ggplot2)#untuk visualisasi data dalam bentuk grafik  library(plotly)#untuk visualisasi grafik pada web  library(tm)#text mining  library(memoise)#session  library(wordcloud)  library(wordcloud2)  library(reshape2)#untuk membentuk data  ``` |

**Listing 3.1** Deklarasi *library* yang digunakan

|  |
| --- |
| ```  hitung\_ulasan = function()  {  wclothing %>%  nrow()  }  hitung\_sentimen = function(x)  {  wclothing %>%  unnest\_tokens(word, `Review Text`) %>%  anti\_join(stop\_words) %>%  inner\_join(get\_sentiments("bing")) %>%  count(sentiment) %>%  filter(sentiment==x)  }  ``` |

**Listing 3.2** Penghitungan jumlah data dan data *cleaning*

|  |
| --- |
| ```  table\_wclothing = function()  {  wclothing %>%  mutate(`Positive Feedback Count` = case\_when(  `Rating` <= 3 ~ "Negatif",  TRUE ~ "Positif"  )) %>%  select(Age,Title,`Review Text`,`Positive Feedback Count`,`Class Name`) %>%  head(50)  }  ``` |

**Listing 3.3** Pengklasifikasian data

|  |
| --- |
| ```  table\_wclothing = function()  {  wclothing %>%  mutate(`Positive Feedback Count` = case\_when(  `Rating` <= 3 ~ "Negatif",  TRUE ~ "Positif"  )) %>%  select(Age,Title,`Review Text`,`Positive Feedback Count`,`Class Name`) %>%  head(50)  }  ``` |

**Listing 3.3** Pengklasifikasian data

|  |
| --- |
| \*\*`ui`\*\*  ```{r ui}  ui = fluidPage(  setBackgroundColor("lightgray"),  title = "Analisis Sentimen terhadap Pakaian Wanita di E-Commerce",  headerPanel("Analisis Sentimen terhadap Pakaian Wanita di E-Commerce"),    fluidRow(  column(  4,  h3("Kata Positif"),  h4(strong(textOutput(outputId = "jumlah\_positif"))),  style = "background-color: white ; background-size: cover;"  ),  column(  4,  h3("Kata Negatif"),  h4(strong(textOutput(outputId = "jumlah\_negatif"))),  style = "background-color: white ; background-size: cover;"  ),  column(  4,  h3("Jumlah Ulasan"),  h4(strong(textOutput(outputId = "jumlah\_ulasan"))),  style = "background-color: white ; background-size: cover;"  ),  ),    sidebarLayout(  sidebarPanel(  style = "background-color: lightgray ; background-size: cover;",  selectInput(  inputId = "Class Name",  label = "Varian Pakaian Wanita",  choices = option\_classname,  multiple = TRUE,  selected = option\_classname[[1]]  ),  column(  8,  h3("Hasil Analisis"),  h4(strong(textOutput(outputId = "d\_analisis"))),  style = "background-color: lightgray; background-size: cover;"  )  ),    mainPanel(  style = "background-color: white ; background-size: cover;",  plotlyOutput(outputId = "plot\_freq\_kata", height = "700px"),  br(),  fluidRow(  column(  4,  h3("Kata Positif"),  h4(strong(textOutput(outputId = "jumlahpositif"))),  style="background-color: lightgray ; background-size: cover;"  ),  column(  4,  h3("Kata negatif"),  h4(strong(textOutput(outputId = "jumlahnegatif"))),  style="background-color: lightgray ; background-size: cover;"  ),  column(  4,  h3("jumlah ulasan"),  h4(strong(textOutput(outputId = "jumlahulasan"))),  style="background-color: lightgray ; background-size: cover;"  )  ),  plotlyOutput(outputId = "plot\_kata\_digunakan", height = "700px"),  h3("Word Clouds", align = "center"),  plotOutput(outputId = "plot\_cloud\_kata", height = "1200px"),  h3("Tabel Ulasan"),  tableOutput(outputId = "plot\_ulasan")  )  )  )  ``` |

**Listing 3.4** *User Interface*

|  |
| --- |
| \*\*`server`\*\*  ```{r server}  server = function(input, output, session)  {  plot\_freq\_kata = reactive({  wclothing %>%  group\_by(`Class Name`) %>%  unnest\_tokens(word, `Review Text`) %>%  group\_by(`Class Name`) %>%  anti\_join(stop\_words) %>%  count(word, sort = T) %>%  na.omit() %>%  filter(n>=1000) %>%  ggplot(aes(x=reorder(word,n), y=n, fill=`Class Name`))+  geom\_bar(stat = "identity")+  coord\_flip()+  labs(  x="Kata",  y="Frekuensi",  title = "Grafik Frekuensi Kata"  )+  theme\_linedraw()  })    output$plot\_freq\_kata = renderPlotly({  ggplotly(plot\_freq\_kata())  })    plot\_kata\_digunakan = reactive({  wclothing %>%  filter(`Class Name` %in% input$`Class Name`) %>%  unnest\_tokens(word, `Review Text`) %>%  anti\_join(stop\_words) %>%  inner\_join(get\_sentiments("bing")) %>%  group\_by(sentiment, `Class Name`) %>%  count(word) %>%  top\_n(10) %>%  ggplot(aes(x=reorder(word,n), y=n, fill=`Class Name`))+  geom\_col(show.legend = T)+  coord\_flip()+  facet\_wrap(~sentiment,scales = "free\_y")+  labs(  x="Kata",  y="Frekuensi",  title = "Grafik Penggunaan Kata"  )+  theme\_linedraw()  })    output$plot\_kata\_digunakan = renderPlotly({  ggplotly(plot\_kata\_digunakan())  })    output$plot\_cloud\_kata = renderPlot({  wclothing %>%  filter(`Class Name` %in% input$`Class Name`) %>%  unnest\_tokens(word, `Review Text`) %>%  anti\_join(stop\_words) %>%  inner\_join(get\_sentiments("bing")) %>%  count(word, sentiment) %>%  acast(word~sentiment, value.var = "n", fill = 0) %>%  comparison.cloud(colors = c("#e82a60", "#5ae8a4"), max.words = 250, scale = c(4,1))  })    hitung\_sentimenx = function(p)  {  wclothing %>%  filter(`Class Name` %in% input$`Class Name`)%>%  unnest\_tokens(word, `Review Text`) %>%  anti\_join(stop\_words) %>%  inner\_join(get\_sentiments("bing")) %>%  count(sentiment) %>%  filter(sentiment==p)  }    hitung\_ulasanx = function()  {  wclothing %>%  filter(`Class Name` %in% input$`Class Name`) %>%  nrow()  }    output$jumlah\_positif = renderText({  hitung\_sentimen("positive")$n  })    output$jumlah\_negatif = renderText({  hitung\_sentimen("negative")$n  })    output$jumlah\_ulasan = renderText({  hitung\_ulasan()  })    output$jumlahpositif = renderText({  hitung\_sentimenx("positive")$n  })    output$jumlahnegatif = renderText({  hitung\_sentimenx("negative")$n  })    output$jumlahulasan = renderText({  hitung\_ulasanx()  })    hasilanalisis = function()  {  if (hitung\_sentimenx("positive")$n>hitung\_sentimenx("negative")$n) {  sprintf("Varian %s Cenderung memiliki feedback Positif", input$`Class Name`)  }  else{  sprintf("Varian %s Cenderung memiliki feedback Negatif", input$`Class Name`)  }  }    output$d\_analisis = renderText({  hasilanalisis()  })    output$plot\_ulasan = renderTable({  table\_wclothing()  })  }  ``` |

**Listing 3.5** *Server* untuk memasukan ke *UI*

|  |
| --- |
| \*\*`run-app`\*\*  ```{r run-app}  shinyApp(ui = ui, server = server)  ``` |

**Listing 3.6** *Running program*

# **KESIMPULAN**

Metode *Latent Dirichlet Allocation* (LDA) dapat digunakan dalam proses klasifikasi data dalam bentuk *text mining*. Penggunaan LDA sering digunakan untuk mengatur atau mengklasifikasikan data menurut topik pemodelan. Dengan mengambil keputusan yang tepat tentunya.

Dari hasil analisis sentiment yang telah dilakukan, dapat disimpulkan bahwa masyarakat lebih banyak merespon sentiment positif dibandingkan sentiment negatif. Menurut hasil wordcloud, semakin besar kata di wordcloud, kata tersebut adalah kata yang paling banyak atau sering digunakan dalam komentar.