

Rancang Bangun Engine Sentimen Analisis Tingkat Popularitas Kopi Berdasarkan Pengolahannya Menggunakan Teknologi Big Data

Ni Putu Nirmala Dewi Widhiasih^{a1}, I Made Agus Dwi Suarjaya^{a2}, Anak Agung Ketut Agung Cahyawan Wiranatha^{b3}

^aProgram Studi Teknologi Informasi, Fakultas Teknik, Universitas Udayana

^bBukit Jimbaran, Bali, Indonesia-8036110

e-mail: 1alanirmala1911@gmail.com, 2agussuarjaya@it.unud.ac.id,

3agung.cahyawan@unud.ac.id

Abstrak

Permintaan produk kopi terus meningkat seiring dengan pertumbuhan penduduk. Permintaan terhadap suatu produk dilihat dari persepsi masyarakat mengenai produk tersebut. Persepsi masyarakat dapat dijumpai melalui kicauan di media sosial Twitter. Persepsi tersebut dapat dimanfaatkan untuk menganalisis sentiment terhadap suatu produk. Pendekatan ini disebut Analisis Sentimen dan untuk menganalisis diperlukan Teknologi *Big Data*. Penelitian ini bertujuan untuk membuat Rancang Bangun *Engine* Sentimen Analisis Tingkat Popularitas Kopi Berdasarkan Pengolahannya Menggunakan Teknologi *Big Data* untuk mengetahui sentimen masyarakat di seluruh dunia terhadap olahan minuman kopi melalui media sosial Twitter. Penelitian ini menggunakan proses *collecting* data *tweet* dari Twitter menggunakan API Twitter, *processing* data, klasifikasi sentimen menjadi positif, negatif, dan netral menggunakan metode Naïve Bayes, serta visualisasi data dalam bentuk grafik. Data yang dihasilkan pada penelitian ini berjumlah 1.253.971 data dalam 1 tahun dengan 1800 data latih dan 450 data uji. Hasil evaluasi dengan klasifikasi Naïve Bayes yaitu *accuracy* 92.67%, *precision* 98.52%, *recall* 88.67%, dan *f-measure* 93.33%.

Kata kunci: *Big Data*, Twitter, Kopi, Naïve Bayes, Sentimen.

Abstract

Demand for coffee products continues to increase along with population growth. The demand for a product is seen from the public's perception of the product. Community perceptions can be found through tweets on social media Twitter. This perception can be used to analyze sentiment towards a product. This approach is called Sentiment Analysis and to analyze it requires Big Data Technology. This study aims to design a Sentiment Engine Analysis of the Level of Popularity of Coffee Based on Processing Using Big Data Technology to find out the sentiments of people around the world towards processed coffee drinks through Twitter social media. This study uses the process of collecting tweet data from Twitter using the Twitter API, data processing, classification into positive, negative, and neutral sentiments using the Naïve Bayes method, as well as data visualization in graphical form. The data generated in this study amounted to 1,253,971 data in 1 year with 1800 training data and 450 test data. The evaluation results with the Naïve Bayes classification are 92.67% accuracy, 98.52% precision, 88.67% recall, and 93.33% f-measure.

Keywords : *Big Data*, Twitter, Coffee, Naïve Bayes, Sentiment.

1. Introduction

Kopi merupakan minuman dari biji kopi yang telah melalui proses sangrai yang kemudian dihaluskan menjadi bubuk. Indonesia merupakan negara dengan penghasil kopi terbaik di dunia yang dibuktikan dengan jumlah ekspor yang dilakukan oleh perusahaan

eksportir komoditas kopi. Indonesia di tahun 2003 mengekspor kopi sebesar 5.800 ton (US \$17,9 juta) serta impor kopi sebesar 1.560 ton (US \$3,56 juta) [1].

Jenis olahan kopi yang pertama kali diperkenalkan hanya ada dua jenis yaitu Americano dan black coffee yang merupakan kopi klasik. Proses pembuatan minuman kopi semakin berkembang dari waktu ke waktu. Jenis kopi yang dapat dikembangkan yakni jenis kopi modern. Kopi modern merupakan jenis olahan kopi dengan menggunakan bahan utama espresso kemudian dikembangkan lagi menjadi beberapa jenis kopi seperti cappuccino, latte, macchiato, dan mochaccino [2].

Franz Georg Kolschitzky merupakan seseorang yang membuka kedai kopi pertama kali karena dia merupakan satu - satunya orang yang memahami nilai dari biji kopi. Franz memilah kopi dan memperlembut rasa dari minuman kopi dengan menambahkan susu dan gula sehingga popularitas kopi semakin meningkat [2].

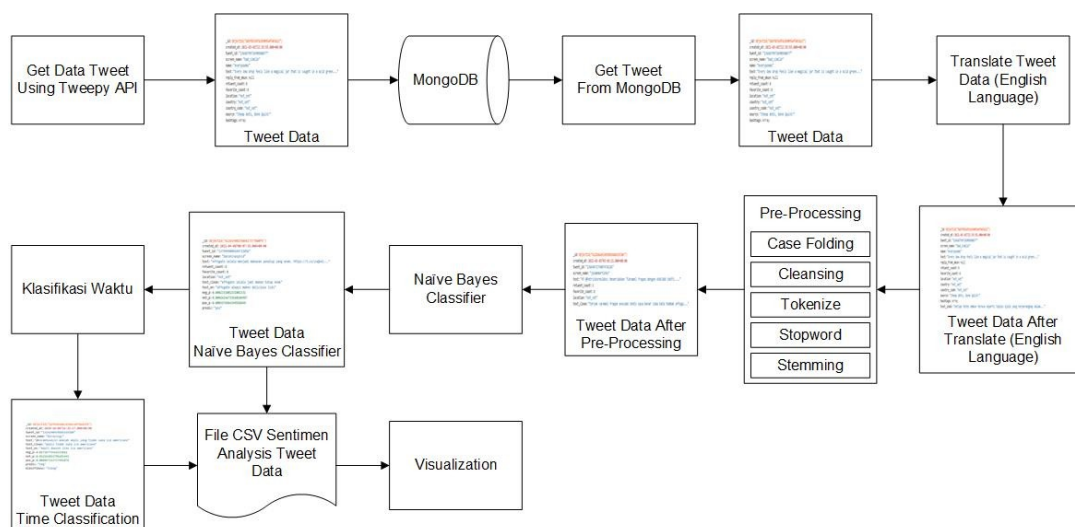
Permintaan terhadap suatu produk dilihat dari presepsi masyarakat mengenai produk tersebut. Masyarakat akan lebih mudah menyampaikan pendapat menggunakan media sosial, salah satunya adalah Twitter. Twitter mempermudah penggunaannya untuk mengetahui kicauan yang paling sering dikicaukan atau populer. Masyarakat dapat memanfaatkan pendapat atau tanggapan dengan menganalisis sentimen, seperti mengekspresikan sentimen positif atau sentimen negatif terhadap suatu produk. Pendekatan ini disebut Analisis Sentimen dan untuk menganalisis diperlukan teknologi big data.

Penelitian sebelumnya mengenai analisis sentimen presepsi kopi organik menggunakan 10 kata kunci yaitu kopi arabika, kopi gayo, kopi tanam sendiri, kopi papua, kopi sakit, kopi hijau, kopi luwak, kopi organik, kopi robusta, dan kopi aceh. Kata kunci tersebut mengambil data menggunakan bahasa Indonesia [3]. Studi lain menggunakan 4 kata kunci yaitu bitter coffee, coffee sick, kopi luwak, dan black coffee [4].

Berdasarkan penelitian tersebut, terdapat dua kebaruan pada penelitian ini. Kebaruan yang pertama menggunakan 10 kata kunci kopi dalam bahasa asing yaitu affogato, americano, blackcoffee, cappuccino, coffeelatte, espresso, flatwhite, macchiato, mochaccino, dan ristretto. Kebaruan yang kedua menggunakan klasifikasi waktu yang bertujuan untuk menganalisis sentimen positif, negatif, dan netral atas kicauan yang diunggah berdasarkan waktu pagi, siang, dan malam. Sentimen olahan minuman kopi dapat digunakan untuk pengambilan keputusan pada bidang wirausaha minuman kopi agar dapat menentukan jenis olahan minuman kopi yang perlu disediakan serta waktu yang tepat dalam menerapkan strategi pemasaran.

2. Metode Penelitian

Metode penelitian adalah alur proses yang digunakan pada penelitian ini digambarkan melalui gambaran umum penelitian.



Gambar 1. Gambaran Umum Penelitian

Gambar 1 menjelaskan serangkaian tahapan yaitu pertama adalah proses *collecting* data menggunakan bahasa pemrograman Python. Proses *collecting* data melalui twitter dengan

menggunakan tweepy API yang akan disimpan pada *database* MongoDB, data yang telah tersimpan pada *database* MongoDB kemudian dilakukan proses *translate* bahasa Inggris karena data yang telah di dapat berupa data dalam berbagai bahasa. Proses *translate* bahasa Inggris dilakukan untuk mempermudah proses selanjutnya yaitu proses *preprocessing*. Proses *preprocessing* meliputi *case folding*, *cleansing*, *tokenize*, *stopword*, dan *stemming*. Proses selanjutnya adalah proses klasifikasi dengan metode Naïve Bayes untuk mengetahui sentimen termasuk ke dalam kelas positif, negatif, dan netral lalu data yang telah diklasifikasi akan tersimpan kembali ke *database* MongoDB. Proses selanjutnya adalah proses klasifikasi waktu untuk mengetahui sentiment yang di posting termasuk ke dalam waktu pagi, siang, atau malam hari. Tahap terakhir yaitu proses visualisasi dalam bentuk grafik menggunakan *tools* Tableau.

3. Kajian Pustaka

Kajian pustaka adalah teori penunjang atau sebagai acuan untuk membuat penelitian yang dijabarkan dibawah ini.

3.1 Analisis Sentimen

Sentimen adalah pendapat atau pandangan terhadap sesuatu berdasarkan perasaan atau emosi seseorang. Analisis sentimen atau *opinion mining* merupakan analisis untuk mengelompokkan polaritas suatu teks pada kalimat menjadi positif, negatif, dan netral. Proses analisis sentimen dapat dilakukan dengan menggunakan pendekatan machine learning [5].

3.2 Pre-Processing Data

Penelitian ini menggunakan lima tahapan pre-processing data yaitu *case folding*, *cleansing*, *tokenizing*, *stopword*, dan *stemming* [6]. Kelima proses tersebut adalah sebagai berikut.

- Case folding digunakan untuk menyelaraskan semua huruf menjadi *lowercase*.
- Cleansing data adalah proses untuk membersihkan data yang tidak diperlukan seperti menghapus link URL, *hashtag*, *mention*, *retweet*, *emoticon*, angka, simbol, tanda baca, karakter selain huruf yang disebut dengan *noisy text*.
- Tokenizing adalah proses untuk memisah kalimat menjadi per-kata berdasarkan spasi yang menyusun kalimat dan tanda baca merupakan bagian dari pemisah kata. Proses ini bertujuan untuk menghapus *whitespace*, spasi berlebih, *enter*, dan tabulasi.
- Stopword merupakan proses untuk menghapus atau meghilangkan kata yang tidak diperlukan karena tidak memiliki makna yang terkandung di dalamnya.
- Stemming bertujuan untuk mengubah kata yang awalnya memiliki imbuhan sehingga menjadi kata dasar.

3.3 Naïve Bayes Classifier

Model klasifikasi yang menggunakan teori keputusan bayes memanfaatkan data latih untuk menghitung kemungkinan setiap kelas berdasarkan nilai feature di dalamnya. Menurut Rifkie Primartha (2021;342) rumus dari Naïve Bayes *Classifier* adalah sebagai berikut [7].

$$P(A|B) = \frac{P(B|A) \times P(A)}{P(B)}$$

$P(A|B)$ = probabilitas A ketika B terjadi atau yang disebut dengan *posterior probability*, $P(B|A)$ probabilitas B ketika A terjadi atau yang disebut dengan *likelihood*, dan $P(B)$ adalah probabilitas (B) atau yang disebut dengan prior *probability*, berlaku ketentuan yaitu $P(B) = 0$ [7].

4. Hasil dan Pembahasan

Hasil analisis dari penelitian ini berdasarkan proses gambaran umum penelitian yang di paparkan sebagai berikut.

4.1. Hasil Pengumpulan Data

Data Twitter yang berhasil dikumpulkan melalui proses *crawling* data Twitter menggunakan Twitter API dengan jumlah 1,253,971 data dalam 1 tahun terhitung bulan Okt 2020 – Okt 2021 yang kemudian data tersebut disimpan kedalam *database* MongoDB.

4.2. Hasil Pre-Processing Data

Proses ini mengubah data kotor menjadi lebih sederhana atau terstruktur agar mudah dilakukannya analisis. *Pre-processing* Data dilakukan dengan beberapa proses yang dijelaskan sebagai berikut

Table 1. Hasil Pre-Processing Data

Proses	Teks	Hasil
Case Folding	RT @bean_inside: Affogato is truly something special. #Coffee #affogato #caffeine https://t.co/5A7Eltx7xW	RT @bean_inside: affogato is truly something special. #Coffee #affogato #caffeine https://t.co/5A7Eltx7xW
Cleansing Data	RT @bean_inside: affogato is truly something special. #Coffee #affogato #caffeine https://t.co/5A7Eltx7xW	affogato is truly something special
Tokenizing	affogato is truly something special	'affogato', 'is', 'truly', 'something', 'special'
Stopword	'affogato', 'is', 'truly', 'something', 'special'	affogato truly something special
Stemming	affogato truly something special	affogato truly some special

4.3 Hasil Pelabelan Data Training

Pelabelan data training pada penelitian ini dilakukan secara manual dengan jumlah data training yang dibuat sebanyak 1800 data. Data training dibuat seimbang dengan 600 data label positif, 600 data label negatif, dan 600 data label netral.

Table 2. Hasil Labeling Data Training

Label	Tweet
pos	flatwhite dolce gusto very cheap very good
pos	delicious fla affogato al caffe vanilla gelato over espresso glass very good
pos	love Italian affogato coffee cart outside
neg	macchiato ovaltine koi really bad
neg	broken barista frap affogato first time bad taste
neg	americano caramel latte large late eat tremor nausea
net	grande withe mocha triple ristretto shot pump vanilla
net	blonde ristretto ice coffee add classic shot white mocha wheat milk
net	belong plain ice americano three shot espresso little sugar

4.4 Hasil Uji Akurasi

Hasil uji akurasi data *training* dibuat dalam bentuk tabel untuk menganalisa nilai *accuracy*, *precision*, *recall*, dan *f1-score*. Tujuan nilai – nilai tersebut untuk mengetahui kualitas model data yang digunakan pada saat klasifikasi.

Table 3. Hasil Pengujian Akurasi

Jumlah	Training	Testing	Akurasi	Presisi	Recall	F1-score
1800	80%	20%	92.67%	98.52%	88.67%	93.33%
1800	70%	30%	91.84%	96.79%	93.41%	95.07%

4.5 Hasil Analisa Metode Naïve Bayes

Hasil analisa metode naïve bayes merupakan proses menentukan tweet masuk ke dalam sentiment negatif, positif, dan netral berdasarkan probabilitas kata pada data training.

```

_id: ObjectId("606e68c572d3eabb84f1b288")
created_at: 2021-04-08T00:07:35.000+00:00
tweet_id: "1379949005694713856"
screen_name: "baconispice"
retweet_count: 0
favorite_count: 0
location: "not_set"
text: "Affogato is always a good dessert. https://t.co/yAQk42Rnc7"
text_en: "Affogato is always a good dessert. https://t.co/yAQk42Rnc7"
text_clean: "affogato always good dessert"
neg_p: 0.00010039152695512498
net_p: 0.00024322023592362885
pos_p: 0.0005099439061703213
predic: "pos"

```

Gambar 2. Hasil Klasifikasi Naïve Bayes

Gambar 2 merupakan hasil analisa metode naïve bayes dengan menentukan setiap tweet masuk ke dalam kelas positif, negatif, atau netral berdasarkan nilai probabilitas tertinggi.

4.6 Hasil Klasifikasi Waktu

Klasifikasi waktu adalah tahap untuk mengelompokan atau menentukan sentiment berdasarkan waktu yang telah di tentukan. Klasifikasi waktu yang digunakan dibagi menjadi tiga yaitu kelas pagi, siang, dan malam. Tabel waktu dan hasil klasifikasi waktu dijelaskan sebagai berikut.

Table 4. Klasifikasi Waktu

Jenis Waktu	Pukul
Pagi	06:00:00 - 11:01:00
Siang	11:00:00 - 17:01:00
Malam	17:00:00 - 22:01:00

```

_id: ObjectId("6386c1dc30aa656cfe5d0013")
created_at: 2021-03-03 06:56:52.000+00:00
tweet_id: "1367006039262994432"
screen_name: "evetnhorizon"
text: "Me, drinking Americano with 6 shots every day: https://t.co/NzrNit9KMO"
text_en: "Me, drinking Americano with 6 shots every day: https://t.co/NzrNit9KMO"
text_clean: "drink americano shot everi day"
neg_p: 0.004517618712980624
net_p: 0.0020673720053508452
pos_p: 0.005303416624171341
predic: "pos"
klasifikasi: "Pagi"

```

Gambar 3. Klasifikasi Waktu Pagi

```
_id: ObjectId("6386c1dd30aa656cfe5d0342")
created_at: 2021-03-01 13:11:44.000+00:00
tweet_id: "1366375605311205386"
screen_name: "KINGKARDED"
text: "I ordered an iced americano today because I hate myself."
text_en: "I ordered an iced americano today because I hate myself."
text_clean: "order ice americano today hate"
neg_p: 0.0019074390121473748
net_p: 0.00012161011796181443
pos_p: 0.00010198878123406426
predic: "neg"
klasifikasi: "Siang"
```

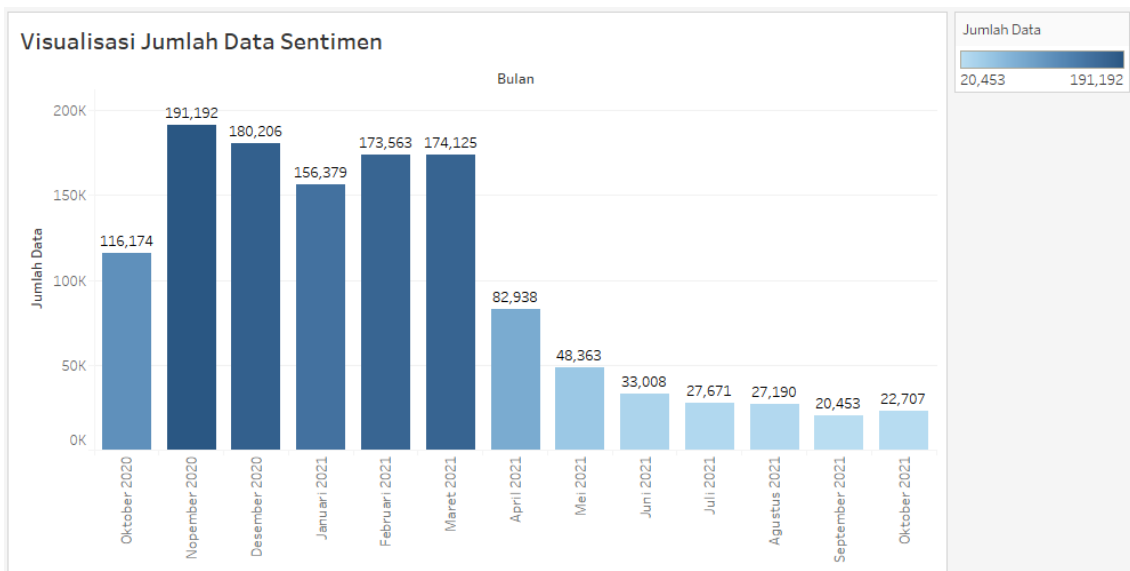
Gambar 4. Klasifikasi Waktu Siang

```
_id: ObjectId("6386c1dd30aa656cfe5d0299")
created_at: 2021-03-01 17:45:01.000+00:00
tweet_id: "1366444377309077506"
screen_name: "itsmemaeghan"
text: "@mahmerc Of the standard drinks the americano is delish very similar i..."
text_en: "@mahmerc Of the standard drinks the americano is delish very similar i..."
text_clean: "standard drink americano delish similar style new irish cream"
neg_p: 0.0008031322156409999
net_p: 0.0066885564878997935
pos_p: 0.0056093829678735335
predic: "net"
klasifikasi: "Malam"
```

Gambar 5. Klasifikasi Waktu Malam

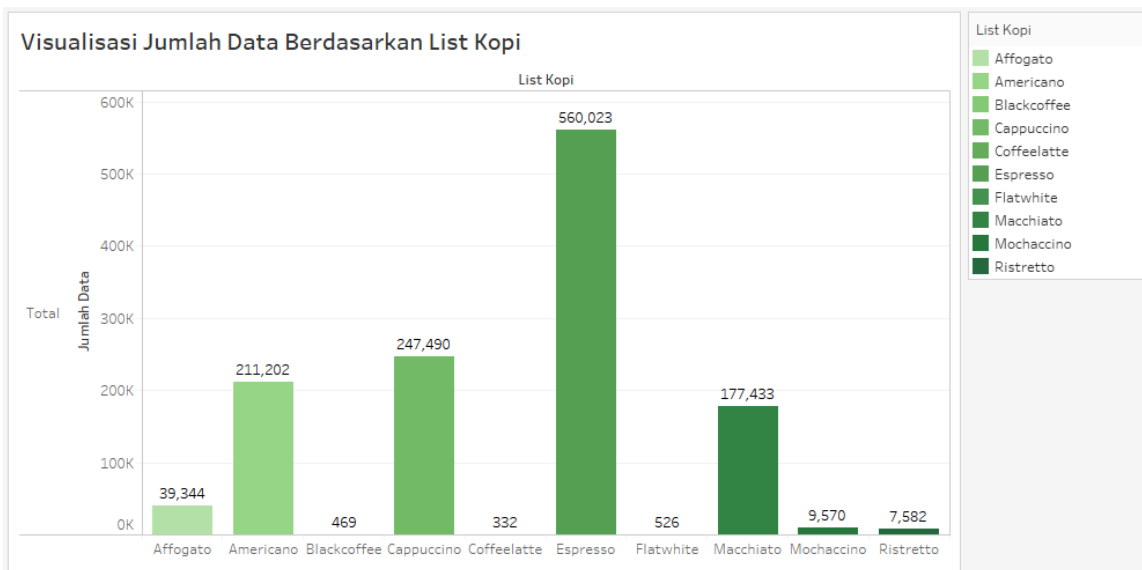
4.7 Hasil Visualisasi

Visualisasi merupakan hasil dari gambaran data sentiment tingkat popularitas kopi berdasarkan pengolahannya yang telah di analisis sebelumnya yang kemudian akan di tuangkan ke dalam bentuk diagram atau grafik. Visualisasi data yang digambarkan merupakan data dalam 1 tahun.



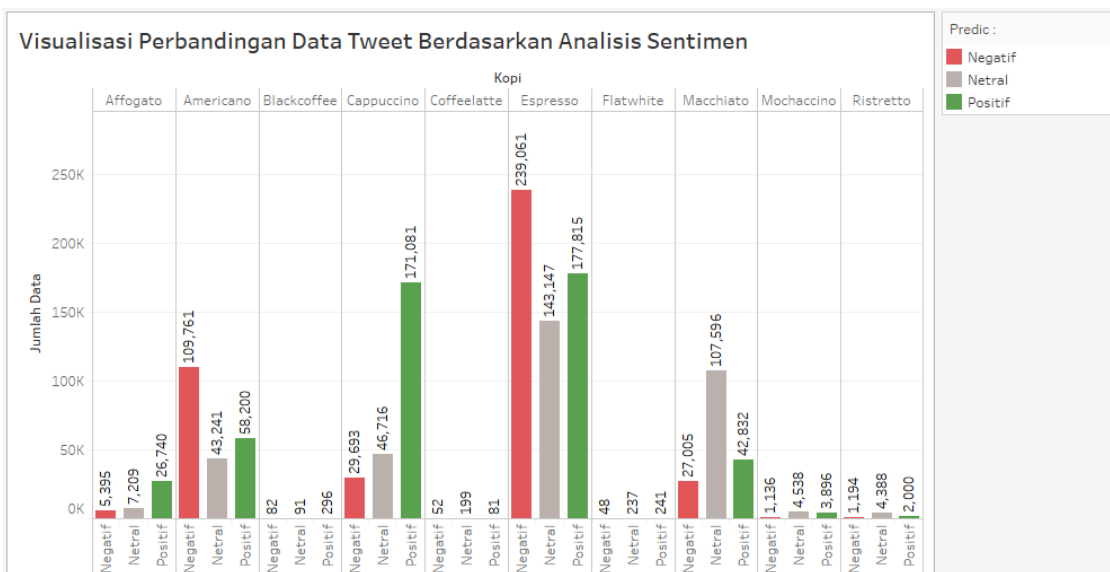
Gambar 6 Jumlah Data Sentimen

Gambar 6 menunjukkan data mengenai olahan minuman kopi internasional dengan total keseluruhan data berjumlah 1.253.969 Data.



Gambar 7 Jumlah Data Berdasarkan List Kopi

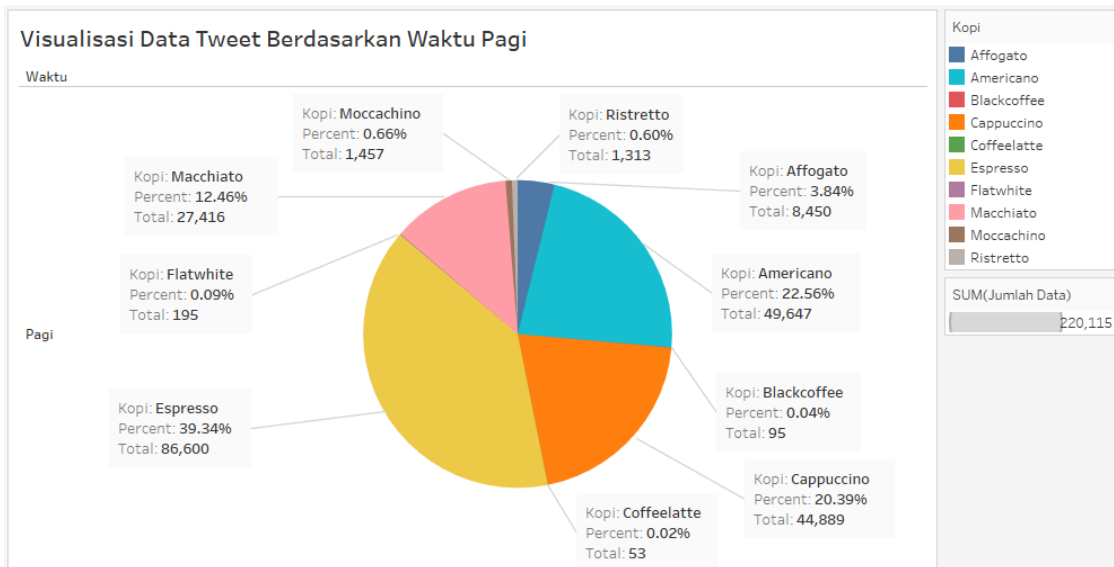
Gambar 7 menunjukkan data mengenai 10 *list* olahan minuman kopi internasional dengan data terbanyak dari 10 *list* olahan minuman kopi yaitu espresso 560.023, urutan kedua yaitu cappuccino 247.490, dan urutan ketiga yaitu americano 211.202. Hasil dari visualisasi diatas dapat disimpulkan bahwa kopi espresso memiliki intensitas tertinggi yang ramai dibicarakan dalam kurun waktu 1 tahun.



Gambar 8 Perbandingan Data Tweet Berdasarkan Analisis Sentimen

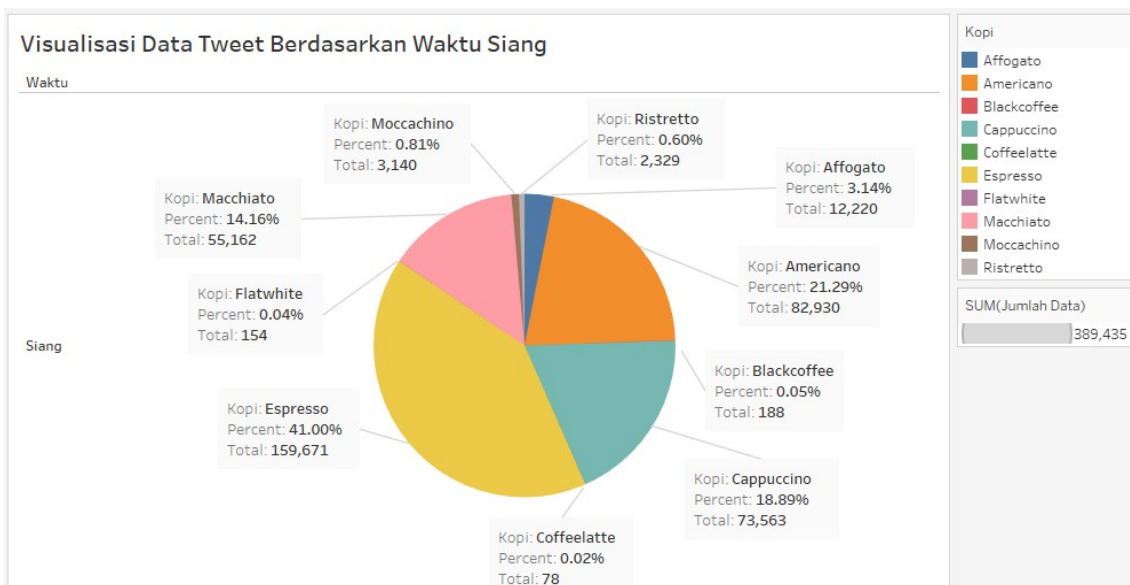
Gambar 8 merupakan perbandingan data *tweet* berdasarkan analisis sentimen 10 *list* olahan minuman kopi dalam kurun waktu 1 tahun. Tingkat popularitas tertinggi dari 10 *list* olahan minuman kopi pada kelas negatif yaitu espresso 239.061, urutan kedua yaitu americano 109.761, urutan ketiga yaitu cappuccino 29.693, urutan keempat yaitu macchiato 27.005, dan urutan kelima yaitu affogato 5.395. Tingkat popularitas tertinggi dari 10 *list* olahan minuman kopi pada kelas positif yaitu espresso 177.815, urutan kedua yaitu cappuccino 171.081, urutan ketiga yaitu americano 58.200, urutan keempat yaitu macchiato 42.832, dan urutan kelima yaitu affogato 26.740. Jenis olahan kopi espresso dan americano memiliki tingkat popularitas tertinggi pada sentimen negatif karena memiliki rasa yang pahit. Efek yang dirasakan oleh orang yang tidak biasa meminum kopi espresso atau americano adalah jantung berdetak

kencang, tidak bisa tidur, mual, pusing, dan sakit kepala. Kopi espresso dan americano memiliki tingkat popularitas tertinggi pada sentimen positif karena kopi espresso ramai dibicarakan sebagai bahan dasar campuran dari jenis kopi yang lain. Kopi espresso dan americano diminum untuk menghilangkan rasa kantuk saat bekerja. Jenis olahan kopi cappuccino mempunyai rasa pahit yang pas dengan campuran susu didalamnya serta memiliki tampilan estetik sehingga menimbulkan daya tarik untuk mengambil foto dan mengunggah di sosial media, hal ini membuat kopi cappuccino memiliki tingkat popularitas tertinggi pada sentimen positif. Kopi macchiato memiliki tingkat popularitas tertinggi pada sentimen positif karena kopi macchiato merupakan minuman dengan tambahan *whipe cream* dan *caramel*. Jenis olahan kopi affogato memiliki tingkat popularitas tertinggi pada sentiment positif karena kopi affogato sering dinikmati sebagai *dessert* atau makanan penutup.



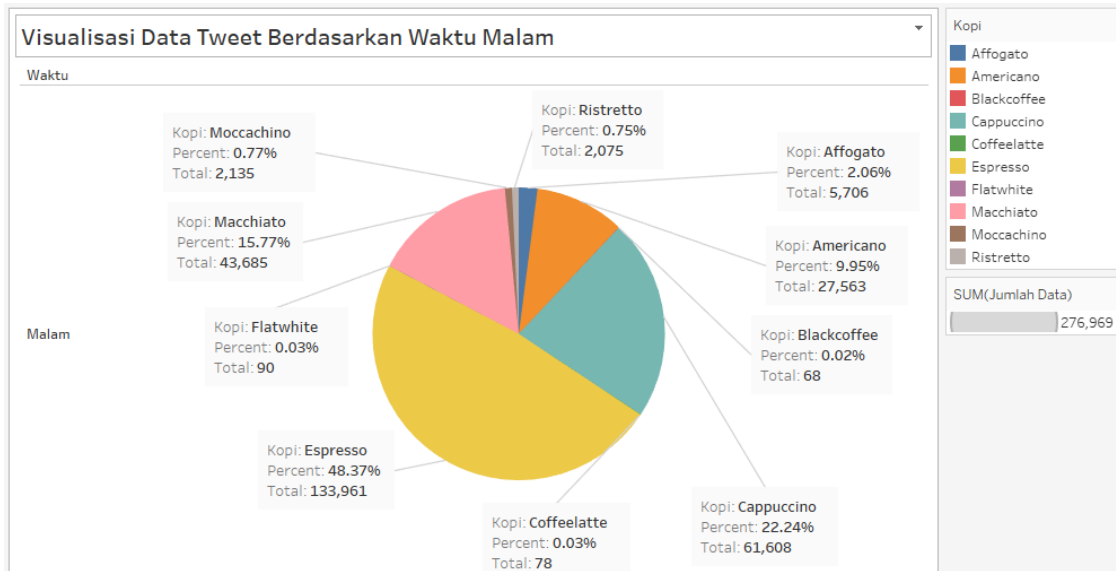
Gambar 9 Data Tweet Berdasarkan Waktu Pagi

Gambar 9 menunjukkan data mengenai 10 *list* olahan minuman kopi internasional dengan data terbanyak di pagi hari yaitu espresso 86.600, urutan kedua yaitu americano 69.647, urutan ketiga yaitu cappuccino 44.889, urutan keempat yaitu macchiato 27.416, dan urutan kelima yaitu affogato 8.450. Visualisasi data tweet berdasarkan waktu pagi berjumlah sebesar 220.115 data.



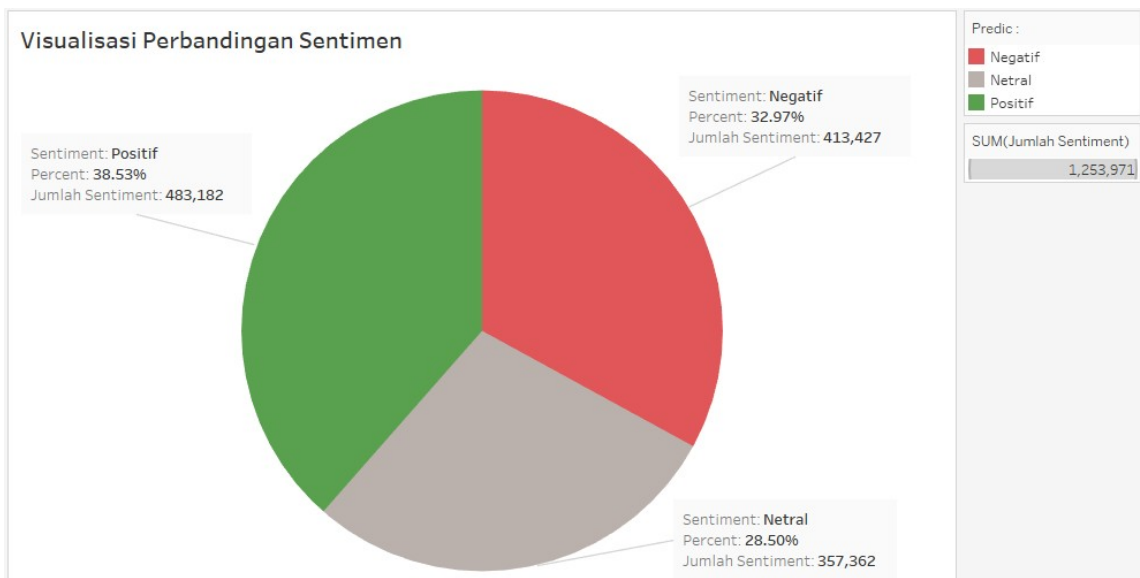
Gambar 10 Data Tweet Berdasarkan Waktu Siang

Gambar 10 menunjukkan data mengenai 10 *list* olahan minuman kopi internasional dengan data terbanyak di siang hari yaitu espresso 159.671, urutan kedua yaitu americano 82.930, urutan ketiga cappuccino 73.563, urutan keempat yaitu macchiato 55.162, dan urutan kelima yaitu affogato 12.220. Visualisasi data tweet berdasarkan waktu siang berjumlah sebesar 389.435 data.



Gambar 11 Data Tweet Berdasarkan Waktu Malam

Gambar 11 menunjukkan data mengenai 10 *list* olahan minuman kopi internasional dengan data terbanyak di malam hari yaitu espresso 133.961 data, urutan kedua yaitu cappuccino 61.608, urutan ketiga yaitu macchiato 43.685, urutan keempat yaitu americano 27.563, urutan kelima yaitu affogato 5.706. Visualisasi data tweet berdasarkan waktu malam berjumlah sebesar 276.969 data.



Gambar 12 Diagram Perbandingan Sentimen

Gambar 12 merupakan diagram perbandingan sentimen kelas negatif, positif, dan netral terhadap 10 *list* olahan minuman kopi Internasional. Kelas Positif mendominasi dengan

persentase tertinggi sebesar 38,53%, kelas Negatif dengan persentase 32,97%, dan Netral sebesar 28,50%.

5. Kesimpulan

Hasil klasifikasi data *tweet* media sosial Twitter terhadap popularitas kopi berdasarkan pengolahannya di seluruh dunia menunjukkan bahwa klasifikasi sentiment positif mendominasi dengan persentase tertinggi sebesar 38,53%, kelas negatif dengan persentase 32,97%, dan netral sebesar 28,50%. Penelitian ini menggunakan metode Naïve Bayes *Classifier* yang menghasilkan akurasi tertinggi sebesar 92,67% menggunakan *training* data sejumlah 1.800 data. Hasil klasifikasi data *tweet* berdasarkan waktu dan jenis menunjukkan bahwa konsumen paling banyak menikmati kopi pada waktu siang hari. Lima jenis kopi yang memiliki persentase tertinggi berdasarkan klasifikasi waktu adalah kopi espresso, cappuccino, americano, macchiato, dan affogato. Lima kopi lainnya yang memiliki persentase terendah diantaranya mochaccino, ristretto, blackcoffee, flatwhite, dan coffeelatte. Jika dari pihak coffee shop ingin meningkatkan minat konsumen untuk mengkonsumsi kelima kopi yang memiliki presentase terendah tersebut dapat menerapkan pemberian promosi pada waktu pagi dan malam hari berupa beli satu gratis satu, beli dua gratis satu, potongan harga pada pembelian kedua, dan paket dengan harga terjangkau.

References

- [1] E. Panggabean, *Buku Pintar Kopi*. Jakarta Selatan: PT AgroMedia Pustaka, 2019.
 - [2] D. H. & A. A. Sastra, *A to Z Memulai dan Mengelola Usaha Kedai Kopi*. Jakarta Selatan: AgroMedia Pustaka, 2020.
 - [3] I. Nuritha, A. A. Arifiyanti, and V. P. Widartha, "Analysis of Public Perception on Organic Coffee through Text Mining Approach using Naïve Bayes Classifier," *Proc. - 2nd East Indones. Conf. Comput. Inf. Technol. Internet Things Ind. EIconCIT 2018*, pp. 153–158, 2018, doi: 10.1109/EIconCIT.2018.8878572.
 - [4] N. Qomariah, "Sentiment Analysis on Coffee Consumer Perceptions on Social Media Twitter Using Multinomial Naïve Bayes," *J. Intell. Comput. Heal. ...*, vol. 2, no. 1, 2021, [Online]. Available: <https://jurnal.unimus.ac.id/index.php/ICHI/article/view/7241>.
 - [5] E. M. Sipayung, H. Maharani, and I. Zefanya, "Perancangan Sistem Analisis Sentimen Komentar Pelanggan Menggunakan Metode Naive Bayes Classifier," *J. Sist. Inf. UNSRI*, vol. 8, no. 1, pp. 958–965, 2016.
 - [6] F. Nurhuda, S. Widya Sihwi, and A. Doewes, "Analisis Sentimen Masyarakat terhadap Calon Presiden Indonesia 2014 berdasarkan Opini dari Twitter Menggunakan Metode Naive Bayes Classifier," *J. Teknol. Inf. ITSmart*, vol. 2, no. 2, p. 35, 2016, doi: 10.20961/its.v2i2.630.
 - [7] D. Kurniawan, *Pengenalan Machine Learning dengan Python*. Jakarta: PT Elex Media Komputindo, 2020.
-