

Identifikasi Faktor Kenaikan Harga Komoditas Pasar Melalui Agregasi Berita Online (Studi Kasus: Jawa Timur)

Rengga Asmara¹, Ferry Astika Sapura², Muhammad Rizal Fauzy³

^{1,2,3} Departemen Teknik Informatika dan Komputer,

^{1,2,3} Politeknik Elektronika Negeri Surabaya

^{1,2,3} Kampus PENS, Jalan Raya ITS, Keputih, Sukolilo, Surabaya

Telp. (031) 594 7280, Fax. (031) 594 6114

e-mail: ¹rengga@pens.ac.id, ²ferryas@pens.ac.id, ³rizalfauzy@it.student.pens.ac.id

Abstrak

Penelitian ini bertujuan untuk membuat sistem monitoring harga komoditas yang mampu mengidentifikasi permasalahan penyebab kenaikan harga komoditas melalui agregasi berita online. Sistem ini memiliki keluaran utama, yaitu monitoring fluktuasi harga pasar dan identifikasi penyebab kasus kenaikan harga komoditas menggunakan data harga komoditas harian yang dikumpulkan dari seluruh pasar induk di Jawa Timur pada Desember 2018 - Mei 2019 yang disediakan oleh SISKAPERBAPO. Untuk mendapatkan hasil identifikasi, sistem melakukan beberapa tahap. Pertama, sistem melakukan labelling terhadap harga pasar yang mengalami kenaikan tidak normal. Selanjutnya, sistem mencari berita terkait kenaikan harga tersebut dan melakukan information retrieval berdasarkan klasifikasi teks terhadap 5W1H pada label why. Hasil dari penelitian ini yaitu, eksperimen pada identifikasi faktor kenaikan harga komoditas menunjukkan bahwa sebesar 1.91% berita memiliki kesesuaian berdasarkan tanggal berita, akurasi berdasarkan relevansi berita dan kasus sebesar 70.49%, dan ekstraksi informasi untuk identifikasi faktor menghasilkan sebesar 39.87% berita relevan menunjukkan hasil identifikasi faktor kenaikan harga komoditas.

Kata kunci: Monitoring System, Tim Pengendalian Inflasi Daerah (TPID), Information Retrieval, Klasifikasi 5W1H

Abstract

This research aims to create a commodity price monitoring system that is able to identify the problems causing the increase in commodity prices through online news aggregation.. This system has main outputs, namely monitoring market price fluctuations and identifying causes of cases of rising commodity prices from online news aggregations, which used daily commodity price data collected from all wholesale markets in East Java in December 2018 - May 2019 provided by SISKAPERBAPO. To get the results of identification, the system performs several stages of processing. First, the system of labeling market prices has increased. Furthermore, the system looks for news related to the price increases and conducts information retrieval based on the classification of text against 5W1H on the label why. The results of this research, namely, the experiment on identifying commodity price increase factors shows that 1.91% of news has conformity based on news date, accuracy based on news and case relevance of 70.49%, and information extraction to identify generating factors of 39.87% relevant news shows results identification of factors in rising prices.

Keywords: Monitoring System, Tim Pengendalian Inflasi Daerah (TPID), Information Retrieval, 5W1H Classification

1. Pendahuluan

Kenaikan harga merupakan masalah ekonomi yang sering kali berdampak buruk pada masyarakat, terutama masyarakat kalangan menengah ke bawah. Banyak faktor yang menyebabkan komoditas pasar mengalami kenaikan harga, antara lain kebijakan pemerintah, permasalahan produksi, distribusi, ketidak-seimbangan permintaan dan persediaan, dan berbagai permasalahan lainnya. Mengingat pentingnya komoditas bahan pokok, kenaikan harga secara langsung dapat mempengaruhi kelangsungan hidup manusia baik dari sisi konsumen, produsen, maupun sisi pemerintah.

Berdasarkan data SISKAPERBAPO (Sistem Informasi Ketersediaan dan Perkembangan Harga Bahan Pokok) provinsi Jawa Timur, sebagian besar komoditas mengalami fluktuasi hampir di setiap pasar induk di Jawa Timur pada tahun 2018. Salah satunya yaitu harga bawang yang naik hingga Rp. 10.000/kg pada bulan Maret 2018 [1]. Berdasarkan data tersebut, dapat diketahui satu kasus kenaikan harga pada komoditas bawang. Jika ditelusuri lebih lanjut maka masih banyak lagi kasus – kasus serupa di titik – titik yang berbeda pada provinsi Jawa Timur, sehingga pemerintah dituntut untuk dapat memahami dan mengatasi berbagai kasus – kasus tersebut untuk dapat menjaga stabilitas harga komoditas.

Adapun badan pemerintah bernama TPID (Tim Pengendalian Inflasi Daerah), berperan menjaga stabilitas harga di daerah. Berbagai rekomendasi pengendalian harga yang dihasilkan TPID dinilai sedikit banyak telah membantu pemangku kepentingan di daerah dalam merumuskan kebijakan terkait pengendalian harga. Langkah kerja yang dilakukan oleh Tim Pelaksana TPID yaitu melakukan pemantauan perkembangan inflasi daerah dan mengidentifikasi berbagai permasalahan terkait pengendalian inflasi di level teknis secara rutin. Hasil pemantauan tersebut kemudian disampaikan kepada Komite Kebijakan bersama dengan usulan rencana aksi dan rekomendasi kebijakan. Komite kebijakan kemudian mengambil keputusan dan memberikan arahan serta masukan kepada Tim Pelaksana terkait pelaksanaan tugas Pokjanas TPID dalam rangka mengatasi permasalahan inflasi daerah [2].

Provinsi Jawa Timur saat ini memiliki sistem informasi yang berfungsi untuk menampung informasi harga komoditas dari setiap pasar induk yang ada di provinsi Jawa Timur. Sistem tersebut yaitu SISKAPERBAPO (Sistem Informasi Ketersediaan dan Perkembangan Harga Bahan Pokok). Sistem ini dapat digunakan untuk tugas pemantauan terhadap harga komoditas di tiap – tiap daerah.

Sesuai dengan langkah kerja, TPID melakukan pemantauan pergerakan harga, kasus kenaikan harga, dan inflasi secara manual, yaitu dengan cara melakukan survey langsung di lokasi pasar terkait. Dengan menggunakan cara tersebut, masih memungkinkan apabila petugas melewatkan informasi terhadap adanya kasus kenaikan harga. Pada penanganan terhadap kasus kenaikan harga, petugas juga harus melaksanakan penyidikan terhadap penyebab kenaikan harga komoditas. Mengingat kompleksitas permasalahan harga komoditas yang mungkin terjadi, langkah kerja TPID semakin rumit dan kurang optimal. Hal tersebut dapat berdampak buruk terhadap kontrol stabilisasi harga komoditas. Berdasarkan kelemahan tersebut, diperoleh masalah yang menjadi acuan dalam riset ini yaitu dibutuhkan sebuah sistem yang digunakan untuk membantu langkah kerja TPID Provinsi Jawa Timur.

Riset ini bertujuan untuk mengimplementasikan sistem informasi untuk membantu TPID dalam melakukan kontrol stabilisasi harga komoditas. Sistem yang dibutuhkan yaitu sistem yang mampu melakukan monitoring terhadap fluktuasi harga komoditas di tiap daerah di Provinsi Jawa Timur dan melakukan identifikasi terhadap kasus kenaikan harga pasar melalui agregasi berita online. Kebutuhan tersebut telah disesuaikan dengan langkah kerja TPID dalam melakukan kontrol stabilisasi harga komoditas.

Beberapa peneliti telah melakukan riset untuk mengatasi permasalahan stabilisasi harga komoditas bahan pokok menggunakan beberapa pendekatan. Pendekatan tersebut yaitu membuat suatu sistem informasi harga komoditas yang diterapkan di berbagai daerah dengan beragam platform. Pendekatan tersebut diimplementasikan oleh Ester K dkk (2017) [3], Fachrizal S dkk (2017) [4], Rahman dan Sri Wahyuni (2017) [5], dan Elis Hernawati dkk (2018) [6]. April LH dan Bayu P (2017) menggunakan pendekatan rancang bangun system informasi dan menerapkan Simple Moving Average dalam perhitungan pergerakan closing price. [7]

Selain menerapkan system informasi harga komoditas. Beberapa peneliti juga menerapkan fitur berupa prediksi harga komoditas secara time series. Shabrina N dkk (2016) melakukan implementasi metode Elman Recurrent Neural Network untuk prediksi harga komoditas pertanian. [8] Performansi sistem diukur dengan menggunakan metode Mean Percentage Error (MAPE). Sistem ERNN dengan algoritma Backpropagation terkendala mendapatkan arsitektur yang optimal karena terjebak di lokal optimum. Hasil yang didapatkan dari penelitian ini adalah prediksi harga bawang merah memperoleh akurasi diatas 75% sedangkan prediksi harga cabai merah memperoleh akurasi dibawah 75%. Sedangkan untuk klasifikasi rekomendasi tanam-harga petani, akurasi yang didapatkan untuk bawang merah kurang dari 75%, sedangkan untuk cabai merah lebih dari 75%.

M Arif R (2017) melakukan penelitian prediksi harga bahan pokok nasional menggunakan ARIMA. [9] Hasil eksperimen menunjukkan bahwa model ARIMA yang dihasilkan mampu memprediksi harga dengan tingkat error rata - rata sebesar 2.22%. Ferlando JS dan Benika N (2018) melakukan penelitian membuat sistem pendukung keputusan monitoring dan peramalan harga beras. [10] Peneliti menggunakan analisis metode Jaringan Saraf Tiruan (JST) yang paling sesuai untuk digunakan sebagai metode permalan/subsistem model SPK. Nilai MSE dan MAPE dari hasil pelatihan, pengujian dan validasi berturut - turut adalah 0,00128 dan 3,57%; 0,0319 dan 5,47%; 0,0052 dan 2,51%. Hasil validasi menunjukkan bahwa hasil peramalan yang dihasilkan oleh SPK memiliki tingkat akurasi 90%.

Amin Pujiati (2018) melakukan penelitian tentang pembuatan model proyeksi untuk memberikan peringatan dini terhadap fluktuasi harga komoditas. [11] Hasil riset menunjukkan metode ARIMA dapat melakukan prediksi terhadap harga komoditas yang dipilih oleh peneliti. Berdasarkan nilai Root Means Square Error (RMSE) dan Mean Absolute Perscent Error (MAPE), keempat model ARIMA dapat digunakan sebagai model peringatan dini dari harga komoditas. Nilai R-Squared dari keempat model lebih besar dari 90% juga menunjukkan bahwa model keseluruhan ARIMA memiliki kinerja yang baik untuk memprediksi harga di masa depan, untuk mengantisipasi permintaan pasar yang berfluktuasi. Vivi AF dkk (2019) melakukan penelitian pembuatan aplikasi peramalan harga sembako. [12] Metode yang digunakan dalam meramalkan harga sembako adalah Single Exponential Smoothing. Dari hasil pengujian sistem didapatkan MAPE terkecil pada alfa=0,9 adalah 0,8%. Sehingga tingkat keakurasian sistem sebesar 99,2%.

Dari beberapa studi literatur yang telah dilakukan, secara garis besar, solusi yang diberikan adalah dengan melakukan implementasi sistem informasi harga komoditas, prediksi pada beberapa harga komoditas, dan Sistem Pendukung Keputusan (SPK). Berdasarkan solusi tersebut, belum ada yang melakukan penanganan dan tindak lanjut terhadap kenaikan harga komoditas.

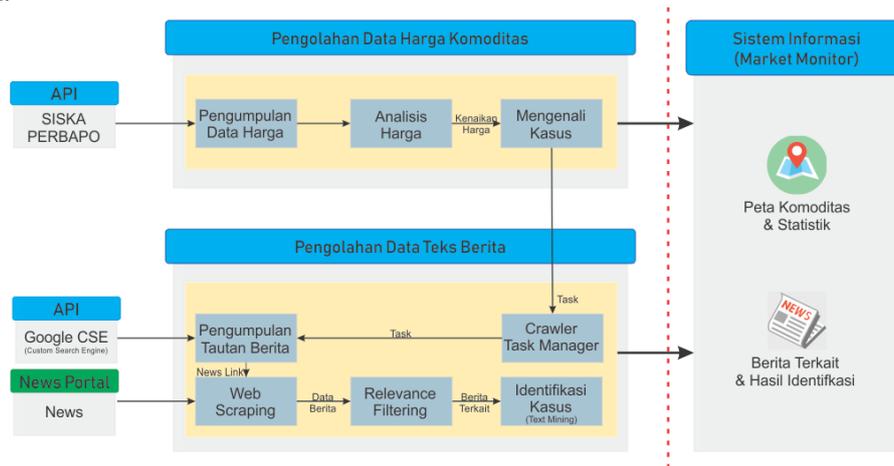
Pada riset ini, sistem baru digunakan untuk membantu TPID dalam melakukan kontrol stabilisasi harga komoditas. Sistem tersebut dapat membantu untuk melakukan identifikasi penyebab kenaikan harga komoditas melalui berita yang beredar di portal berita online. Sistem tersebut mencakup beberapa bagian proses, yaitu (1) pencarian berita berdasarkan kata kunci yang diberikan pada portal berita, (2) penyeleksian berita berdasarkan relevansi berita dan kasus kenaikan harga, (3) ekstraksi informasi dari teks berita yang telah dilakukan seleksi. Pendekatan

pada solusi ini belum pernah diimplementasikan oleh peneliti sebelumnya, sehingga penulis akan mencoba menyelesaikan permasalahan menggunakan pendekatan ini.

2. Metode Penelitian

2.1. Diagram Sistem

Dalam perencanaan sebuah sistem membutuhkan sebuah diagram umum. Gambar 1 merupakan diagram sistem dari riset ini. Secara garis besar, terlihat dari diagram umum bahwa diagram terbagi menjadi 3 sub-sistem, yaitu (1) pengolahan data harga komoditas, (2) pengolahan data teks berita, dan (3) sistem informasi (Market Monitor). Masing – masing sub-sistem memiliki peranan penting dalam sistem yang selanjutnya akan dibahas pada sub bab di bawah ini.



Gambar 1. Diagram sistem

A. Pengolahan Data Harga Komoditas

Dari diagram sistem pada gambar 3, data utama yang dibutuhkan meliputi data harga komoditas yang telah disediakan oleh sistem informasi yang telah tersedia di Provinsi Jawa Timur yaitu SISKAPERBAPO. Daftar komoditas yang disediakan seperti pada tabel 1. Data harga komoditas dikumpulkan pada level pasar induk, per kota/kabupaten dan provinsi. Untuk mendapatkan harga pada level kota/kabupaten dan provinsi, perlu dilakukan operasi rata – rata harga berdasarkan kota/kabupaten. Hal yang sama juga dilakukan pada data harga pada level provinsi. Semua data tersebut akan dikumpulkan per hari untuk mendapatkan data yang valid dan terbaru.

Tabel 1. Komoditas tersedia

Komoditas			
Beras Bengawan	Beras Mentik	Beras IR 64	Daging Sapi
Daging Ayam Potong (Broiler)	Daging Ayam Kampung	Cabai Merah Biasa	Cabai Keriting

1. Pengumpulan Data Harga Komoditas

Data harga komoditas dikumpulkan dengan cara mengakses API penyedia data secara berkala, sehingga data yang didapatkan merupakan data yang valid dan terbaru. Gambar 2 merupakan *response* akses API yang tersedia. Dari *response* yang didapatkan, selanjutnya data akan dilakukan proses *parsing* atau penyesuaian struktur data pada *response* dan struktur database harga komoditas. Selanjutnya data akan disimpan. Setelah data disimpan, data dapat diolah menjadi 2 bentuk, yaitu perhitungan prediksi inflasi dan pengelompokan kenaikan harga komoditas.

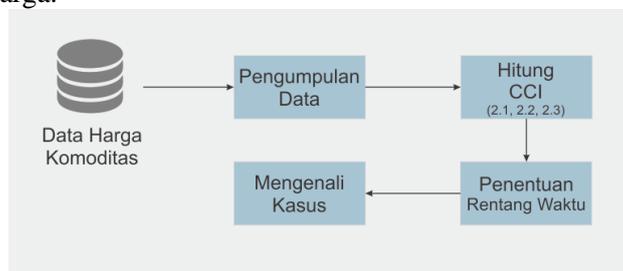
```

{
  "status": 200,
  "status_message": "Data Found 2018-12-26",
  "result": [
    {
      "id": "2434944",
      "commodity_id": "2",
      "commodity_name": "Beras Bengawan",
      "commodity_title": "Beras Bengawan / kg",
      "harga": "12500",
      "data_tgl": "2018-12-26",
      "update_at": "2018-12-27 00:13:30",
      "update_ke": "",
      "naik": "0",
      "turun": "0",
      "id_pasar": "1",
      "nama_pasar": "Pasar Tambahrejo",
      "kota": "Kota Surabaya"
    },
    ...
  ]
}
    
```

Gambar 2. Respon API SISKAPERBAPO

2. Analisis Harga

Pada tahap ini dilakukan pengenalan mengenai harga yang memiliki peningkatan yang dianggap tidak normal. Gambar 3 menunjukkan langkah – langkah yang diperlukan untuk melakukan analisis harga.



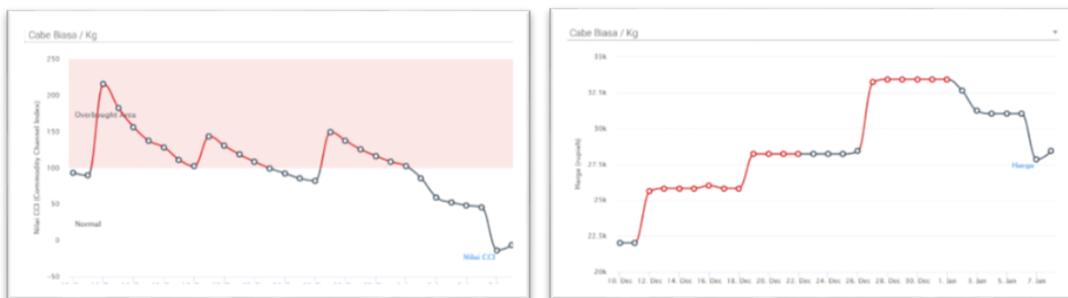
Gambar 3. Diagram analisis harga komoditas

Berdasarkan gambar 3 diperlukan perhitungan nilai indeks CCI (Commodity Channel Index) dengan menerapkan persamaan (1), (2), dan (2). [13] Setiap harga pada komoditas dan pasar induk tertentu akan melalui tahap tersebut secara time series.

$$\bar{X} = \frac{i}{N} \sum_{i=1}^N x_i \quad (1)$$

$$MD = \frac{i}{N} \sum_{i=1}^N |x_i - \bar{X}| \quad (2)$$

$$CCI = \frac{(x_i - \bar{X})}{0.015 MD} \quad (3)$$



(a)

(b)

Gambar 4. (a) Grafik Commodity Channel Index (b) Grafik time series harga dan kasus

Dari gambar grafik pada gambar 4 (a) menunjukkan bahwa komoditas tersebut mengalami kenaikan yang melebihi batas normal. Hal tersebut terlihat dari nilai indeks CCI yang melampaui 100 atau telah memasuki *overbought area* sehingga harga dikategorikan terlampaui tinggi.

3. Mengenali Kasus Kenaikan Harga Komoditas

Pada tahap ini dilakukan pengenalan terhadap kasus kenaikan harga. Kasus tersebut didapatkan dari rentang waktu dimana indeks harga terdapat pada *overbought area* seperti pada gambar 4 (b). Grafik tersebut merupakan grafik harga komoditas dan kasus kenaikan harga. Terlihat bahwa terdapat 2 kasus yang dapat dikenali ditunjukkan oleh garis berwarna merah. Kasus yang terjadi dapat dikenali karena pada rentang waktu tersebut harga komoditas telah memasuki *overbought area* pada grafik CCI. Selanjutnya kasus kenaikan harga komoditas akan disimpan dengan atribut jenis komoditas, wilayah, waktu awal kenaikan harga, dan waktu akhir kenaikan harga. Proses pengenalan terhadap kasus kenaikan harga dilakukan pada level kota/kabupaten dan provinsi. Selanjutnya, kasus kenaikan harga komoditas ini digunakan sebagai pembangkit task untuk pengolahan data teks berita.

B. Pengolahan Data Teks Berita

Paragraf harus teratur. Semua paragraf harus rata, yaitu sama-sama rata kiri dan dan rata kanan. Pada blok pengolahan data teks berita akan dilakukan operasi data meliputi pengumpulan data berita terkait dengan kasus kenaikan harga yang terjadi dan pemrosesannya hingga mendapatkan informasi yang berguna.

1. News Crawler Task Manager

News Crawler Task Manager adalah sebuah sistem yang mengatur semua task terhadap penelusuran dan pengumpulan data berita terkait. Sistem ini bertugas untuk melakukan penelusuran berita terhadap portal berita online dan melakukan pengumpulan data teks berita. Proses tersebut tergolong kompleks. Tingginya kasus kenaikan harga yang terjadi pada berbagai daerah Jawa Timur dapat mengakibatkan adanya penambahan beban terhadap sistem pengumpul berita. Karena hal tersebut, sistem ini dibutuhkan untuk mengantisipasi *task* untuk mencari berita dalam jumlah besar.

Sistem akan menyelesaikan tugas secara berkala sesuai dengan task yang diberikan. Teknologi yang dipilih pada pengembangan sistem ini yaitu NodeJS karena memiliki kelebihan *asynchronous*, cepat, dan ringan. Sistem tersebut juga dilengkapi dengan konsep *thread pool* untuk mengatur task - task yang masuk untuk diselesaikan. *thread pool* akan membangkitkan *thread worker* sesuai dengan *task* yang masuk pada sistem ini dan akan menyesuaikan dengan jumlah maksimum *worker* yang telah diberikan. Hal tersebut akan sangat memudahkan kontrol terhadap komponen sistem ini dan juga dapat menyesuaikan dengan kebutuhan server atau *environment requirements*.

2. Pengumpulan Tautan Berita

Pada tahap ini melakukan pengumpulan tautan atau URL berita terkait dengan metode crawling. Crawling adalah proses penelusuran yang dilakukan ke internet. Pada metode crawling ini memanfaatkan teknologi yang telah tersedia yaitu Google Custom Search Engine API (CSE). Google CSE merupakan sebuah layanan pencarian situs berbasis API. Layanan ini akan mengembalikan tautan dari situs hasil dari pencarian berdasarkan kata kunci yang diminta.

Sebelum melakukan pencarian, sistem terlebih dahulu menyiapkan sebuah kata kunci yang digunakan untuk mencari berita terkait dengan kasus kenaikan harga komoditas pasar. Kata kunci tersebut harus mengandung kata yang spesifik terhadap jenis komoditas serta titik dimana kasus kenaikan harga terjadi. Berikut merupakan format kata kunci yang digunakan pada sistem.

“Harga [kata kunci komoditas] [kata kunci titik wilayah]”

Dari hasil tahap, didapatkan daftar hasil pencarian yang mengandung tautan berita. Tautan berita tersebut selanjutnya akan diakses seluruhnya untuk mendapatkan teks berita pada tahap selanjutnya.

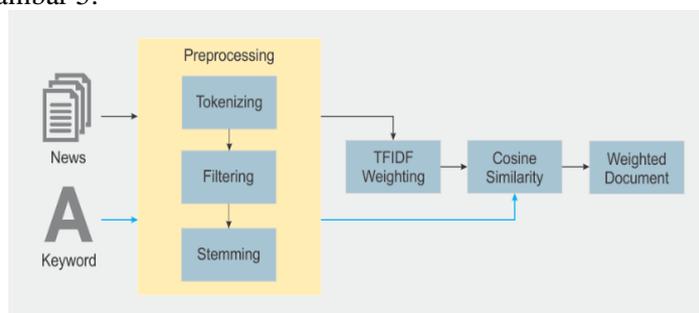
3. Web Scraping

Tahap web scraping merupakan tahap melakukan ekstraksi informasi dari sebuah situs yang memuat artikel berita terkait berdasarkan tag HTML. Sebelum melakukan hal tersebut diperlukan data portal berita yang dituju beserta dengan tag HTML dari informasi yang akan diambil. Portal berita yang ditarget yakni portal berita yang kredibel dan memberikan informasi mengenai harga komoditas pasar. Portal berita yang dikumpulkan sejumlah 130 portal berita.

4. Relevance Filtering Berita

Hasil *crawling* dan *scraping* pada tahap sebelumnya masih memiliki hasil yang tidak sesuai dengan kasus. Diantaranya terdapat link portal berita yang tidak menunjukkan artikel atau teks, link portal berita dengan kasus yang tidak sesuai dengan komoditas dan titik kasus, tanggal berita yang tidak sesuai dengan tanggal kasus, serta permasalahan lain. Oleh karena itu diperlukan sebuah sistem yang mampu melakukan penyaringan data hasil pengumpulan data berita.

Berita yang sesuai yaitu berita yang memiliki kecocokan berdasarkan tanggal berita, jenis komoditas yang bermasalah, dan wilayah tempat terjadi permasalahan tersebut. Sehingga sistem penyaringan dibuat berdasarkan parameter tersebut. Pada tahap pertama, berita diseleksi berdasarkan tanggal berita dan tanggal kasus yang terjadi. Selanjutnya untuk melakukan seleksi berdasarkan komoditas dan wilayah, diterapkan metode *similarity* pada kata kunci pencarian dan teks berita, atau disebut juga dengan *keyword similarity*. Pada tahap seleksi berita ini, diterapkan metode untuk mengukur similarity dari kata kunci pencarian dan masing – masing berita yang didapatkan dari proses pengumpulan berita sebelumnya. Konsep metode ini dijelaskan pada gambar 5.

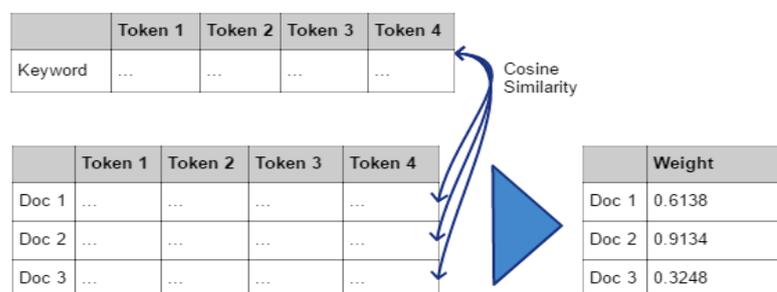


Gambar 5. Diagram *keyword similarity*

Seperti pada gambar 5 menunjukkan diagram proses komponen filtering berita. Pada diagram tersebut menunjukkan bahwa terdapat beberapa tahap. Pada tahap *preprocessing* terdapat *tokenizing*, *filtering*, dan *stemming*. *Tokenizing* merupakan tahap pengubahan data berupa teks berita menjadi kata – kata. *Filtering* merupakan tahap untuk menghilangkan kata – kata umum atau disebut juga sebagai *stopword*. *Stemming* merupakan pemisahan kata terhadap imbuhan yang melekat pada kata tersebut, sehingga yang didapatkan adalah kata dasar. Berikut merupakan contoh hasil dari *preprocessing* pada teks data berita.

- Teks berita : “NGANJUKTIMES, MALANG – Menghadapi Ramadhan dan Idul Fitri, Satgas Pangan Kabupaten Malang melakukan rakor, Senin (30/04) di ruang Rupatama Polres Malang.”
- Hasil *preprocessing* : ['nganjuktimes', 'malang', 'menghadapi', 'ramadhan', 'idul', 'fitri', 'satgas', 'pangan', 'kabupaten', 'malang', 'rakor', 'senin', 'ruang', 'rupatama', 'polres', 'malang']

Selanjutnya, setelah melalui proses *preprocessing*, data berita akan melalui proses *TFIDF Weighting*. Pada proses ini, akan dilakukan perhitungan pembobotan menggunakan metode TFIDF untuk masing - masing data berita. Hasil dari proses ini yaitu sebuah matriks kata terhadap dokumen berita beserta nilai TFIDF-nya. Setelah mendapatkan nilai *TFIDF Weighting* dari dokumen berita. Selanjutnya dihitung nilai cosine similarity [14] masing – masing berita terhadap kata kunci. Nilai yang dihitung adalah nilai TFIDF dokumen berita dengan nilai TFIDF kata kunci. Gambar 6 adalah gambar ilustrasi dari proses perhitungan cosine similarity.

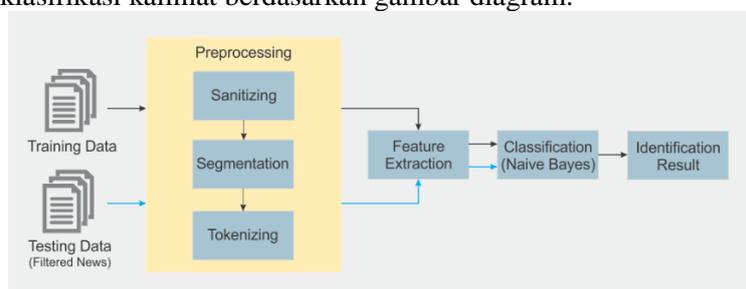


Gambar 6. Perhitungan *cosine similarity* pada *keyword*

Langkah akhir hasil cosine dapat difilter menggunakan *threshold* yang telah ditentukan sehingga komponen ini dapat memisahkan berita yang relevan untuk dilakukan proses lebih lanjut.

5. Identifikasi Faktor Kenaikan Harga

Setelah mendapatkan berita yang relevan, selanjutnya dapat dilakukan proses identifikasi menggunakan metode ekstraksi informasi berdasarkan 5W + 1H [15]. Pada tahap ini, berita akan dilakukan pengklasifikasian berdasarkan elemen 5W + 1H kalimat. Gambar 7 merupakan diagram proses klasifikasi kalimat berdasarkan gambar diagram.



Gambar 7. Diagram Klasifikasi 5W+1H

Gambar diagram komponen diatas menunjukkan urutan proses identifikasi berita. Sebelum melakukan pengolahan data, disiapkan terlebih dahulu *Training Data* berupa 100 berita terkait kenaikan harga komoditas yang telah dikumpulkan secara manual. Data tersebut selanjutnya dipisah per kalimat dan dilakukan pemberian label secara manual untuk setiap kalimatnya. Label yang diberikan adalah (*what, who, when, where, why, how, other*).

Setelah didapatkan *Training Data*, maka baik data *training* maupun *testing* akan melewati tahap *preprocessing*. *Sanitizing* adalah tahap untuk menghilangkan data yang tidak diperlukan seperti tag HTML hasil scraping yang masih tersisa. *Segmentation* adalah tahap untuk membagi teks dokumen berita menjadi kalimat – kalimat yang nantinya akan diproses tersendiri. Dalam proses ini dokumen akan dipisahkan per titik. Namun, tentu terdapat masalah jika terdapat penggunaan titik untuk singkatan, gelar, pemisah angka, dll. Maka diperlukan beberapa *rule* untuk mengatasi hal tersebut. Pertama, pemecahan kalimat berdasarkan titik dilakukan dengan

pemberian label terhadap kalimat, yaitu berupa *parent* (induk kalimat) dan *fraction* (bagian kalimat). Kemudian untuk setiap kalimat akan dilakukan pengecekan terhadap *rule* berikut :

- Beri label *parent* untuk semua kalimat.
- Jika kata terakhir pada kalimat berupa digit seperti “000” dan awal kalimat selanjutnya berupa digit, maka beri label *fraction* untuk kalimat selanjutnya.
- Jika kata terakhir pada kalimat berupa gelar/singkatan awalan dan awal kalimat selanjutnya berupa gelar/singkatan awalan, maka beri label *fraction* untuk kalimat selanjutnya.
- Jika kata pertama pada kalimat tidak dalam bentuk kapital, maka beri label *fraction* untuk kalimat ini.

Tokenizing adalah tahap untuk memisahkan kalimat berdasarkan spasi dan penghapusan tanda baca, sehingga yang didapatkan adalah murni sebuah kata.

Setelah melalui tahap *Preprocessing*, masing – masing kalimat akan melalui tahap ekstraksi fitur berdasarkan 12 fitur yang tertera pada tabel 2. Masing – masing fitur memiliki syarat untuk menentukan apakah data kalimat berita tersebut memenuhi syarat tersebut atau tidak.

Tabel 2. Fitur klasifikasi 5W+1H

No	Fitur	Aturan
1	Ada Angka	Terdapat angka pada kalimat.
2	Ada Bulan	Terdapat kata nama bulan ('januari', 'februari', ..., 'january', 'february', ..., 'jan', 'feb', ...)
3	Ada Penunjuk	Terdapat kata preposisi penunjuk ('di', 'pada', 'dalam', 'atas', 'antara', 'dari', 'ke', 'kepada', 'akan', 'terhadap').
4	Ada Penghubung	Terdapat kata sambung / konjungsi ('sebab', 'karena', 'akibat', 'hingga', 'sehingga', 'lantaran').
5	Ada Kapital	Terdapat kata yang diawali huruf kapital pada huruf pertamanya di tengah kalimat.
6	Ada Jabatan	Terdapat kata yang menjelaskan jabatan ('menteri', 'presiden', 'wakil', 'kepala', 'gubernur', ...).
7	Ada Lokasi	Terdapat kata yang terkait dengan nama tempat ('negara', 'provinsi', 'pulau', 'kota', 'kabupaten', ...).
8	Ada Proses	Terdapat kata yang menunjukkan sebuah proses ('lalu', 'kemudian', 'selanjutnya', 'akan').
9	Ada Satuan	Terdapat kata mengenai satuan/unit ('mg', 'g', 'ons', 'kg', 'miligram', ...).
10	Ada Proyek	Terdapat kata yang mendeskripsikan konten sebuah pengembangan pada berita ('perumahan', 'rumah', 'rusun', 'apartemen', 'mess', ...).
11	Ada Masalah	Terdapat kata yang menjelaskan penyebab dari sebuah masalah pada berita ('bebas', 'macet', 'demo', 'mangkrak', 'sendat', ...).
12	Muncul di Awal	Kalimat bertempat pada awal paragraf.

Berikut merupakan contoh hasil ekstraksi fitur berdasarkan syarat pada tabel fitur 3.4.

Kalimat : “Jombang, Jurnaljatim.com Menjelang perayaan Natal dan Tahun Baru 2018, sekitar satu minggu ini, harga sejumlah kebutuhan pokok di sejumlah pasar tradisional di Kabupaten Jombang mulai merangkak naik”

Hasil :

Tabel 3. Hasil Klarifikasi Berdasakan 12 Fitur 3W+1H

	F1	F2	F3	F4	F5	F6	F7	F8	F9	F10	F11	F12
1	0	1	0	1	0	1	0	0	1	1	1	

Selanjutnya, berdasarkan 12 fitur tersebut akan dilakukan klasifikasi menggunakan metode *Naïve Bayes single label* untuk menentukan apakah kalimat tersebut terklasifikasi untuk salah satu fitur (*what, who, when, where, why, how, other*). Kemudian, kalimat yang memiliki label *Why* menunjukkan bahwa teridentifikasi penyebab kenaikan harga yang dimuat pada berita tersebut.

C. Sistem Informasi (Market Monitor)

Market monitor adalah sebuah aplikasi sistem informasi yang dibuat untuk menampilkan seluruh hasil pengolahan data harga pada riset ini. Sistem informasi akan menampilkan informasi dasar seperti harga terkini, tabel harga, hingga statistik. Sistem informasi ini memiliki fitur utama yaitu, (1) Peta Komoditas dan Statistik, Merupakan fitur yang menampilkan harga masing – masing komoditas dan wilayah dengan memanfaatkan Sistem Informasi Geografis (GIS). Penerapan GIS dianggap lebih mudah untuk dipahami oleh pengguna karena data dapat langsung dipetakan sesuai dengan titik lokasinya. Sedangkan statistik menampilkan data menggunakan grafik berbasis *time series* memudahkan pengguna memahami pertumbuhan harga komoditas pasar pada kurun waktu tertentu. (2) Berita Terkait dan Identifikasi Kasus, Merupakan fitur yang membantu identifikasi penyebab kenaikan harga komoditas pasar melalui berita terkait. Sistem akan menyajikan berita yang relevan terhadap masing – masing kasus dan menunjukkan hasil identifikasi kasus yang sesuai. Fitur ini terbagi menjadi level per kota dan provinsi.

3. Hasil dan Pembahasan

Bab ini menjelaskan mengenai uji coba sistem dan analisa hasil uji coba. Uji coba yang dibahas yaitu uji coba berdasarkan keluaran sistem yaitu sistem informasi monitoring harga komoditas dan hasil identifikasi faktor kenaikan harga komoditas. Pengujian dan analisa keluaran - keluaran tersebut digunakan sebagai tolok ukur keberhasilan sistem pada penelitian ini.

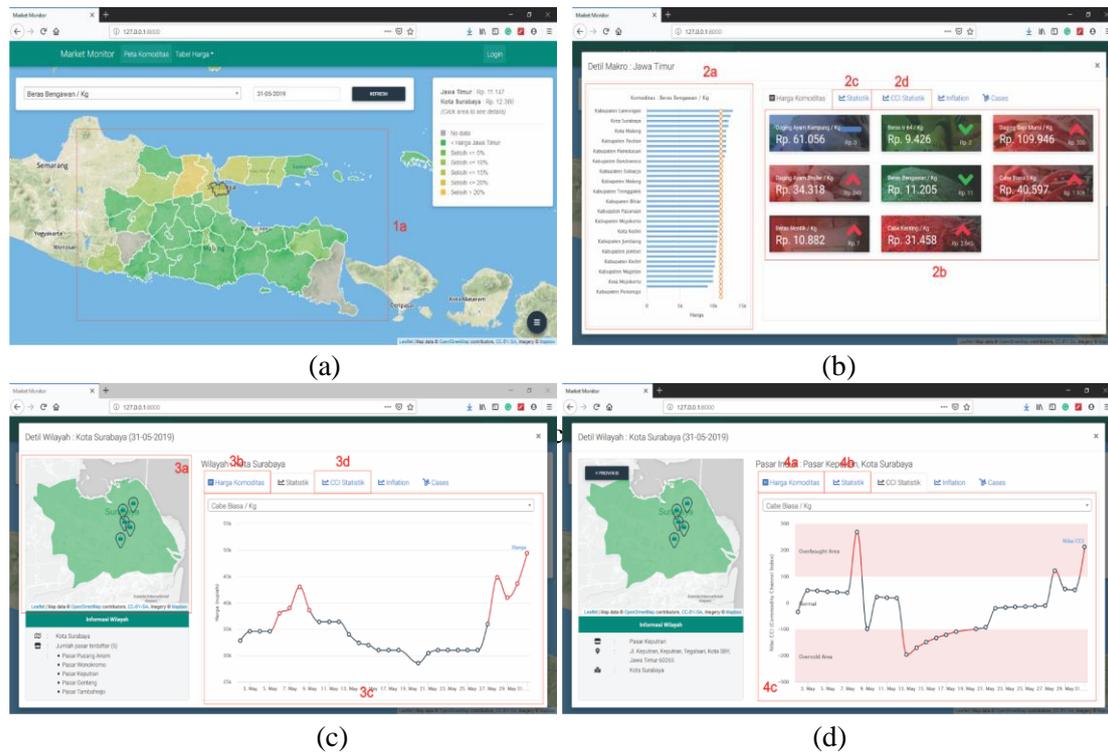
A. Monitoring Harga Komoditas Pasar (Market Monitor)

Sistem informasi *Market Monitor* digunakan untuk memvisualisasikan harga komoditas dalam beberapa bentuk, yaitu berupa data, grafik, dan peta. Oleh karena itu dibutuhkan beberapa skenario uji coba berdasarkan fitur dan ekspektasi keluaran/tampilan pada setiap fiturnya. Tabel 3 menunjukkan skenario uji coba fitur, ekspektasi keluaran, beserta hasil uji coba.

Tabel 4. Hasil uji coba fitur market monitor

Fitur	Ekspektasi Keluaran	Hasil
1. Peta komoditas	a. Peta <i>Choropleth</i> provinsi Jawa Timur berdasarkan harga komoditas	Berhasil
2. Detil makro (Jawa Timur)	a. Grafik pemeringkatan harga komoditas kota/kabupaten	Berhasil
	b. Harga komoditas provinsi	Berhasil
	c. Statistik <i>time series</i> harga komoditas & kasus	Berhasil
	d. Statistik <i>time series CCI Index</i>	Berhasil
3. Detil wilayah kota/kabupaten	a. Peta lokasi pasar induk	Berhasil
	b. Harga komoditas kota/kabupaten	Berhasil
	c. Statistik <i>time series</i> harga komoditas & kasus	Berhasil
	d. Statistik <i>time series CCI Index</i>	Berhasil
4. Detil pasar induk	a. Harga komoditas pasar induk	Berhasil
	b. Statistik <i>time series</i> harga komoditas & kasus	Berhasil
	c. Statistik <i>time series CCI Index</i>	Berhasil

Gambar 8 (a) – 8 (d) merupakan tangkapan layar dari sistem informasi *Market Monitor* dan korelasi keluaran berdasarkan tabel 4.



Gambar 8. (a) Tangkapan layar peta komoditas (b) Tangkapan layar detil macro (c) Tangkapan layar detil wilayah (d) Tangkapan layar detil pasar induk

Market Monitor berperan untuk menampilkan seluruh informasi hasil dan keluaran dari penelitian ini. Salah satu fungsinya yaitu untuk memonitoring harga komoditas. Data harga diolah dan ditampilkan dalam bentuk informasi, grafik, dan peta. Berdasarkan tabel hasil uji coba sistem informasi *Market Monitor*, hasil menunjukkan bahwa pada tiap - tiap fitur dan keluaran berhasil memvisualisasikan data harga komoditas. Sistem menampilkan harga dan kasus kenaikan harga yang terjadi pada lingkup makro, kota/kabupaten dan pasar induk. Oleh karena itu, sistem informasi ini dapat melakukan monitoring harga komoditas dan kasus kenaikan harga yang terjadi.

B. Identifikasi Faktor Kenaikan Harga Komoditas Melalui Berita Online

Pada sub-bab ini dibahas mengenai uji coba dan analisa keluaran dari fitur identifikasi faktor kenaikan harga komoditas melalui berita *online*. Fitur ini merupakan fitur untuk menampilkan hasil pencarian penyebab kasus kenaikan harga yang terjadi. Hasil pencarian yang dibutuhkan meliputi berita terkait kenaikan harga dan penyebab kenaikan harga. Hasil tersebut harus sesuai dengan waktu terjadinya kasus kenaikan harga, jenis komoditas yang terkait, dan wilayah terjadinya kenaikan harga.

Eksperimen yang dilakukan untuk menguji hasil identifikasi faktor kenaikan harga komoditas dilakukan berdasarkan 2 proses. Eksperimen dilakukan terhadap pencarian berita relevan dan ekstraksi fitur. Eksperimen akan dilakukan terhadap kasus yang terjadi pada bulan Desember 2018, April 2019, dan Mei 2019. Pada eksperimen ini, sistem akan dijalankan untuk melakukan pencarian berita secara keseluruhan. Sistem akan dijalankan untuk melakukan proses pencarian berita relevan dan ekstraksi fitur. Sebagai data pembanding, peneliti juga melakukan

proses pencarian berita relevan dan ekstraksi fitur secara manual. Selanjutnya dari kedua hasil tersebut, hasil dapat dibandingkan untuk mendapatkan pengukuran kinerja fitur ini.

Percobaan pada eksperimen fitur ini dilakukan terhadap bulan Desember 2018, April 2019, dan Mei 2019. Masing – masing bulan dipilih karena memiliki karakteristik kenaikan harga karena *event* yang terjadi pada bulan tersebut. Tabel 4 adalah beberapa hasil percobaan beserta penjelasannya.

Tabel 5. Hasil eksperimen identifikasi faktor kenaikan harga komoditas

Desember 2018			
Komoditas :	Daging sapi	Wilayah :	Kab. Jombang
Rentang Waktu Kasus :	15 Des 2018 – 24 Des 2018		
Hasil			
URL : https://faktualnews.co/2018/12/22/harga-daging-di-jombang-mahal-polisi-melakukan-penyelidikan/114290/			
Headline : Harga Daging di Jombang Mahal, Polisi Melakukan Penyelidikan (22 Desember 2018)			
<ol style="list-style-type: none"> 1. Lebih lanjut, dia menegaskan bahwa tidak segan - segan akan memberikan tindakan tegas jika mendapati unsur kesengajaan hingga praktik penimbunan sapi yang berdampak pada mahalnya harga daging sapi ini 2. Para pedagang tidak mengetahui pasti penyebab mahalnya harga daging ini 3. Hanya saja mereka menduga ini disebabkan oleh naiknya harga pakan ternak ayam 4. Sedangkan untuk sapi, para penjual daging ini menduga karena sulitnya pihak jagal mendapatkan pasokan sapi hidup dari peternak 			
April 2019			
Komoditas :	Cabai keriting	Wilayah :	Prov. Jawa Timur
Rentang Waktu Kasus :	6 Apr 2019 – 25 April 2019		
Hasil			
URL : http://infopublik.id/kategori/nusantara/342959/harga-cabai-rawit-di-pasar-jatim-rata-rata-rp19-000-per-kg			
Headline : Harga Cabai Rawit di Pasar Jatim Rata-Rata Rp19.000 Per Kg (18 April 2019)			
<ol style="list-style-type: none"> 1. "Menurut data harga cabai dari kelompok sentra petani cabai Kabupaten Kediri dan sekitarnya harga cabai naik turun tetapi masih dalam batas kewajaran Khususnya cabai merah besar biasa dan varitas gada serta imola harganya terus mernajak naik, karena petikan setiap hari berkurang," ujar Suyono 2. Meskipun cuaca saat ini sebgaiian besar daerah di Jawa Timur curah hujan tinggi tetapi tidak berpengaruh terhadap daserah sentra sayur - mayur tetap bisa memanen hasil tanamannya meskipun mulai berkurang, sehingga harga sayur mayur termasuk cabai stoknya msih cukup 			
Mei 2019			
Komoditas :	Ayam potong	Wilayah :	Kota Mojokerto
Rentang Waktu Kasus :	3 Mei 2019 – 11 Mei 2019		
Hasil			
URL : https://news.detik.com/berita-jawa-timur/d-4542524/harga-bumbu-dapur-dan-ayam-di-kota-mojokerto-kembali-naik-ini-penyebabnya			
Headline : Harga Bumbu Dapur dan Ayam di Kota Mojokerto Kembali Naik, Ini Penyebabnya (9 Mei 2019)			

1. "Harganya naik karena suplier juga menaikkan harga ayam potong hidup.Saya pun terpaksa ikut menaikkan harga supaya tak rugi," ungkap perempuan berhijab ini
2. "Awal puasa biasanya petani libur ke kebun sehingga panen berkurang.Biasanya kenaikan ini tak lama akan turun kembali," tegasnya
3. Tersendatnya pasokan menjadi salah satu penyebabnya

Setelah dijalankan beberapa percobaan pada sistem ini, didapatkan beberapa nilai probabilitas dan akurasi, yaitu dari probabilitas berita yang terseleksi berdasarkan tanggal, akurasi metode seleksi berita menggunakan text mining, dan probabilitas hasil identifikasi faktor dari teks berita online. Tabel 5 menunjukkan hasil eksperimen pada fitur ini.

Tabel 6. Hasil akurasi identifikasi faktor kenaikan harga komoditas

Bulan	Total Berita	Seleksi Tanggal	Seleksi Relevansi	Identifikasi Faktor
Desember 2018	5286	1.21%	68.75%	29.63%
April 2019	3898	1.72%	73.13%	42.42%
Mei 2019	5278	2.8%	69.59%	47.56%
Rata - rata		1.91%	70.49%	39.87%

Berdasarkan tabel hasil eksperimen, terdapat beberapa hasil yang dapat dianalisa. Pertama yaitu berdasarkan kolom seleksi tanggal didapatkan probabilitas berita yang terseleksi berdasarkan tanggal kasus bernilai sangat kecil yaitu 1.91%. Sebagian besar berita memiliki tanggal yang berbeda dari kasus. Kedua, yaitu berdasarkan kolom seleksi relevansi berita menggunakan metode *text mining*, metode *keyword similarity* memberikan hasil sebesar 70.49%. Ketiga, berdasarkan kolom identifikasi faktor mendapatkan probabilitas sebesar 39.87%. Hasil tersebut merupakan hasil ekstraksi fitur 5W+1H. Hasil tersebut ada dikarenakan beberapa penyebab yaitu tidak semua teks berita menuliskan faktor penyebab kenaikan harga yang terjadi, kemungkinan kedua adalah terdapat adanya kesalahan sistem ekstraksi fitur sehingga faktor kenaikan harga tidak dapat teridentifikasi.

Dari beberapa analisa tersebut, dapat disimpulkan bahwa fitur identifikasi kenaikan harga komoditas menggunakan ekstraksi informasi 5W+1H dapat digunakan untuk membantu lembaga pemerintah dalam melakukan identifikasi penyebab kenaikan harga. Meskipun hasil berita terkait dan identifikasi faktor terbilang kecil, namun pada beberapa kasus menunjukkan hasil yang sesuai dengan kondisi yang terjadi di lapangan.

4. Kesimpulan

Berdasarkan hasil eksperimen, dapat disimpulkan bahwa (1) Market Monitor berhasil menunjukkan keluaran visualisasi data harga komoditas pada fitur monitoring harga komoditas. Visualisasi yang dihasilkan berupa informasi harga, grafik time series, dan peta komoditas (2) Fitur identifikasi faktor kenaikan harga komoditas menunjukkan beberapa hasil eksperimen yaitu berita terseleksi berdasarkan tanggal berita sebesar 1.91%, metode seleksi berdasarkan relevansi komoditas dan wilayah kasus menggunakan metode *keyword similarity* dengan akurasi sebesar 70.49%, hasil ekstraksi informasi menunjukkan bahwa 39.87% berita relevan menunjukkan hasil identifikasi faktor kenaikan harga komoditas. Dari beberapa kesimpulan yang dihasilkan per fitur, secara keseluruhan sistem informasi Market Monitor mampu menyelesaikan permasalahan monitoring kasus kenaikan harga komoditas. Dengan adanya system Market Monitor ini, TPID dapat terbantu dalam melakukan kontrol terhadap stabilisasi harga komoditas pasar dengan fitur – fitur yang telah disediakan sistem.

Berdasarkan kesimpulan yang telah dibuat, saran yang dapat digunakan oleh peneliti selanjutnya adalah Peneliti selanjutnya dapat menerapkan metode lain dalam melakukan seleksi

berita online dan ekstraksi informasi berdasarkan 5W+1H sebagai solusi lain maupun peningkatan performa dalam menyelesaikan permasalahan kenaikan harga komoditas.

Daftar Pustaka

- [1] D. Jatim, "SISKAPERBAPO." <http://siskaperbapo.com/harga/grafik> (accessed May 14, 2018).
- [2] BANK INDONESIA, "Pokjanas TPID - Bank Sentral Republik Indonesia." www.bi.go.id/id/moneter/koordinasi-pengendalian-inflasi/tim (accessed May 14, 2018).
- [3] E. Kojongian, H. F. Wowor, and S. D. S. Karouw, "Sistem Informasi Komoditas Pasar di Kota Manado Berbasis Android," *E-Journal Tek. Inform.*, vol. 12, no. 1, 2017.
- [4] F. Sinaga, A. P. Munir, and S. B. Daulay, "PENGEMBANGAN SISTEM INFORMASI DENGAN ANDROID UNTUK HARGA KOMODITAS PERTANIAN (Information System Development with Android for Agriculture Commodity Price)," *J.Rekayasa Pangan dan Pert.*, vol. 5, no. 3, pp. 554–557, 2017.
- [5] Rahman and S. Wahyuni, "Desain Sistem Informasi Harga Pangan Realtime Sebagai Instrumen Kebijakan Pengendalian Inflasi Daerah," *J. INSYPRO (Information Syst. Process.*, vol. 2, no. 2, pp. 1–9, 2017.
- [6] E. Hernawati, E. Rosely, and R. Wulan, "Aplikasi Informasi Harga Kebutuhan Pokok Masyarakat Real Time (Studi Kasus : Diskominfo Kabupaten Bandung)," *J. Sist. dan Teknol. Inf.*, vol. 6, no. 4, pp. 197–203, 2018.
- [7] A. L. Hananto and B. Priyatna, "RANCANG BANGUN APLIKASI INFORMASI HARGA PRODUK," *TechnoXplore J. Ilmu Komput. Teknol. Inf.*, vol. 2, no. 1, pp. 10–20, 2017.
- [8] S. Nanggala, D. Saepudin, and F. Nhita, "ANALISIS DAN IMPLEMENTASI ELMAN RECURRENT NEURAL NETWORK UNTUK PREDIKSI HARGA KOMODITAS PERTANIAN (Analysis and Implementation of Elman Recurrent Neural Network for Predicting Agricultural Commodities Prices)," in *e-Proceeding of Engineering*, 2016, vol. 3, no. 1, pp. 1253–1262.
- [9] M. A. Rasyidi, "Prediksi Harga Bahan Pokok Nasional Jangka Pendek Menggunakan ARIMA," *J. Inf. Syst. Eng. Bus. Intell.*, vol. 3, no. 2, pp. 107–112, 2017.
- [10] F. J. Simanungkalit and B. Naibaho, "Sistem Pendukung Keputusan Monitoring dan Peramalan Harga Beras di Kabupaten Deli Serdang , Sumatera Utara (Decision Support System for Monitoring and Forecasting of Rice Price in Deli Serdang , North Sumatera)," *J. Agritech*, vol. 38, no. 2, pp. 208–216, 2018.
- [11] A. Pujiati, S. Oktavilia, and N. Damayanti, "The Projection Model As an Early Warning of Food Price Commodity Fluctuation," in *International Conference on Economics, Business and Economic Education 2018*, 2018, vol. 2018, pp. 860–868, doi: 10.18502/kss.v3i10.3178.
- [12] V. A. Fitria, R. D. Indahsari, and M. S. Masykur, "Pembuatan Aplikasi Peramalan Harga Sembako Di Kota Malang Berbasis Web," *J. Sist.*, vol. 8, no. 1, pp. 1–9, 2019.
- [13] M. Maitah, P. Prochazka, M. Cermak, and K. Šrédli, "Commodity channel index: Evaluation of trading rule of agricultural commodities," *Int. J. Econ. Financ. Issues*, vol. 6, no. 1, pp. 176–178, 2016.
- [14] B. Herwijayanti, D. E. Ratnawati, and L. Muflikhah, "Klasifikasi Berita Online dengan

- menggunakan Pembobotan TF-IDF dan Cosine Similarity,” *J. Pengemb. Teknol. Inf. dan Ilmu Komput.*, vol. 2, no. 1, pp. 306–312, 2018.
- [15] R. B. Hutama, A. R. Barakbah, and A. Helen, “Indonesian News Auto Summarization in Infrastructure Development Topic using 5W + 1H Consideration,” in *2017 International Electronics Symposium on Knowledge Creation and Intelligent Computing (IES-KCIC)*, 2017, pp. 258–264.



Digital Zone: Jurnal Teknologi Informasi dan Komunikasi is licensed under a [Creative Commons Attribution International \(CC BY-SA 4.0\)](https://creativecommons.org/licenses/by-sa/4.0/)
