# BAB IV HASIL PENELTIAN DAN PEMBAHASAN

Pada bab ini penulis menguraikan hasil dari implementasi dan uji coba sistem. Bab ini terdiri dari pengumpulan data, pengembangan model, implementasi sistem, hasil implementasi, panduan penggunaan aplikasi, rancangan eksperimen, analisis dan hasil eksperimen.

## Pengumpulan Data

Penelitian ini menggunakan data historis nilai tukar mata uang asing terhadap rupiah (kurs), data klimatologi, data kualitas udara. Data-data tersebut dipilih berdasarkan jenis data tersebut yaitu *time series* dengan format waktu “mm/dd/yyyy hh:mm”, selain itu juga digunakan beberapa data dari penelitian sebelumnya, guna untuk membandingkan hasil dari sistem yang akan dikembangkan dengan sistem yang dikembangkan sebelumnya. Maka dari itu, pembahasan lebih detil akan dipaparkan pada sub-bab selanjutnya.

### Data Nilai Tukar Mata Uang Asing

Data diperoleh dari situs Kementrian Perdagangan Republik Indonesia (http://www.kemendag.go.id). Data berisikan nilai tukar mata uang asing terhadap mata uang rupiah atau kurs dengan rentang per bulan, mulai dari 1 Januari 2001 sampai dengan 1 Juli 2018 dengan jumlah data sebanyak 211 baris. Untuk mengakses data yang tersedia, penulis mengunjungi situs Kemenag RI lalu memilih menu statistik yang berada pada bagian bawah situs, lalu penulis memilih menu nilai tukar rupiah dan memilih menu export to excel, atau bisa diakses melalui alamat (www.kemendag.go.id/id/economic-profile/economic-indicators/exchange-rates).

Setelah melakukan serangkaian proses untuk mengakses data, data yang didapatkan berbentuk tabel dengan jumlah baris sebanyak 211 baris, dan jumlah atribut sebanyak 10 atribut, dengan nilai atribut mewakili nilai tukar mata uang tersebut terhadap rupiah. Berikut adalah atribut-atribut yang terdapat pada data tersebut:

* Waktu data
* Dolar Amerika (USD)
* Yen Jepang (JPY)
* Poundsterling Inggris (GBP)
* Franc Swiss (CHF)
* Dolar Singapore (SGD)
* Riggit Malaysia (MYR)
* Dolar Hongkong (HKD)
* Dolar Austrailia (AUD)
* Dolar Canada (CAD)

Untuk lebih lengkapnya data dapat dilihat pada lampiran, namun penulis mengutip beberapa data yang disajikan pada tabel 4.1, seperti yang sudah dijelaskan sebelumnya data berisikan beberapa atribut yang merepresentasikan nilai tukar mata uang asing terhadap rupiah.

Tabel 4. Data Nilai Tukar Mata Uang Asing

|  |  |  |  |  |  |  |
| --- | --- | --- | --- | --- | --- | --- |
| DateTime | USD | JPY | GBP | CHF | ... | CAD |
| 01/01/2001 00:00 | 9.45000 | 8.13149 | 13.81498 | 5.74364 | ... | 6.28451 |
| 02/01/2001 00:00 | 9.83500 | 8.45297 | 14.17963 | 5.85644 | ... | 6.43274 |
| 03/01/2001 00:00 | 10.40000 | 8.37000 | 14.85227 | 6.01400 | ... | 6.60907 |
| 04/01/2001 00:00 | 11.67500 | 9.42066 | 16.74548 | 6.76774 | ... | 7.56988 |
| 05/01/2001 00:00 | 11.05800 | 9.21733 | 15.76486 | 6.21098 | ... | 7.15591 |
| ... | ... | ... | ... | ... | ... | ... |
| 03/01/2018 00:00 | 13.75600 | 12.90553 | 19.36503 | 14.38763 | ... | 10.64665 |
| 04/01/2018 00:00 | 13.87700 | 12.71895 | 19.11904 | 14.04698 | ... | 10.80554 |
| 05/01/2018 00:00 | 13.95100 | 12.83973 | 18.55344 | 14.11117 | ... | 10.83194 |
| 06/01/2018 00:00 | 14.40400 | 13.03707 | 18.83468 | 14.44230 | ... | 10.86439 |
| 07/01/2018 00:00 | 14.41300 | 12.98294 | 18.91203 | 14.59250 | ... | 11.03346 |

### Data Klimatologi

Data ini diperoleh dari stasiun pemantuan milik lembaga meteorologi dan klimatologi Galicia yaitu MeteoGalicia yang berlokasi di kota Mabegondo, Provinsi A Coruna, daerah komUnitas otonom Galicia, Spanyol. Data ini berisikan data klimatologi dengan rentang per hari mulai dari tanggal 06 Juli 2016 sampai dengan 6 Juli 2017 sebanyak 365 baris data. Data ini merupakan data yang digunakan pada penelitian sebelumna, yaitu DWP yang dibahas pada bab 2 sebelumnya. Parameter-parameter yang terdapat pada data ini diantaranya:

Waktu data

Cakupan awan rata-rata per hari dengan satuan % (Persen)

Suhu rata-rata per hari dengan satuan oC (derajat celcius)

Kecepatan angin rata-rata per hari dengan satuan km/h (Kilometer per jam)

Arah angin rata-rata per hari dengan satuan o (derajat)

Curah hujan rata-rata per hari dengan satuan L/m2 (Liter per meter kuadrat)

Parameter tersebut merupakan hasil analisis pada penelitian DWP sebelumnya (Putra dkk., 2017). Untuk lebih lengkapnya data dapat dilihat pada lampiran dan untuk kutipan datanya bisa dilihat pada tabel 4.2.

Tabel 4. Kutipan data klimatologi

|  |  |  |  |  |  |
| --- | --- | --- | --- | --- | --- |
| DateTime | CloudCoverage | Temperature | WindSpeed | WindDirection | Rainfall |
| 07/06/2016 00:00 | 40.8 | 21.3 | 5.47 | 315 | 0 |
| 07/07/2016 00:00 | 20.9 | 20.1 | 6.41 | 315 | 0 |
| 07/08/2016 00:00 | 27.2 | 19.5 | 7.02 | 315 | 0 |
| 07/09/2016 00:00 | 23.2 | 19.1 | 5.94 | 315 | 0 |
| 07/10/2016 00:00 | 77.5 | 18.7 | 5.44 | 180 | 0 |
| ... | ... | ... | ... | ... | ... |
| 07/02/2017 00:00 | 12.6 | 18.9 | 7.34 | 315 | 0 |
| 07/03/2017 00:00 | 13.3 | 21.8 | 4.86 | 315 | 0 |
| 07/04/2017 00:00 | 18.7 | 24 | 6.59 | 225 | 0 |
| 07/05/2017 00:00 | 81.1 | 19.2 | 8.35 | 225 | 0 |
| 07/06/2017 00:00 | 58.3 | 17.9 | 6.48 | 315 | 0 |

### Data Kualitas Udara

Data ini hampir mirip dengan data klimatologi sebelumnya, data ini diperoleh dari stasiun pemantuan milik lembaga meteorologi dan klimatologi Galicia yaitu MeteoGalicia. Data ini berisikan nilai-nilai partikel yang terkandung dalam udara dengan rentang per hari mulai dari tanggal 06 Juli 2016 sampai dengan 6 Juli 2017 sebanyak 365 baris data. Parameter-parameter yang terdapat pada data ini diantaranya:

Data Karbon Monoksida (CO) per hari dengan satuan ppm (Part per Million)

Data Nitrogen Monoksida (NO) per hari dengan satuan ppm (Part per Million)

Data Nitrogen Dioksida (NO2) per hari dengan satuan ppm (Part per Million

Data Nitrogen Oksida (NOX) per hari dengan satuan ppm (Part per Million)

Data Ozone (O3) per hari dengan satuan ppm (Part per Million)

Data Particulate Matter 10 mikronmeter per hari

Data Particulate Matter 25 mikronmeter per hari

Data Sulfur Dioksida (SO2) per hari dengan satuan ppm (Part per Million)

Data ini juga digunakan pada penelitian DWP sebelumnya (Putra dkk., 2017). Tabel 4.3 merupakan kutipan data yang digunakan selama penelitian, untuk 366 baris data secara lengkap dapat dilihat pada lampiran.

Tabel 4. Kutipan data Kualitas Udara

|  |  |  |  |  |  |  |  |  |
| --- | --- | --- | --- | --- | --- | --- | --- | --- |
| DateTime | CO | NO | NO2 | NOX | O3 | PM10 | PM25 | SO2 |
| 07/06/2016 00:00 | 0.13 | 3 | 15 | 19 | 51 | 18 | 10 | 1 |
| 07/07/2016 00:00 | 0.11 | 1 | 10 | 10 | 56 | 14 | 7 | 1 |
| 07/08/2016 00:00 | 0.1 | 1 | 8 | 8 | 59 | 12 | 8 | 1 |
| 07/09/2016 00:00 | 0.1 | 2 | 10 | 11 | 57 | 12 | 7 | 1 |
| 07/10/2016 00:00 | 0.11 | 2 | 10 | 12 | 53 | 12 | 11 | 1 |
| ... | ... | ... | ... | ... | ... | ... | ... | ... |
| 07/02/2017 00:00 | 0.17 | 2 | 35 | 37 | 10 | 14 | 20 | 7 |
| 07/03/2017 00:00 | 0.17 | 23 | 43 | 78 | 1 | 11 | 14 | 9 |
| 07/04/2017 00:00 | 0.17 | 31 | 42 | 90 | 1 | 0 | 14 | 9 |
| 07/05/2017 00:00 | 0.18 | 32 | 41 | 90 | 1 | 0 | 12 | 7 |
| 07/06/2017 00:00 | 0.18 | 32 | 61 | 90 | 1 | 0 | 12 | 7 |

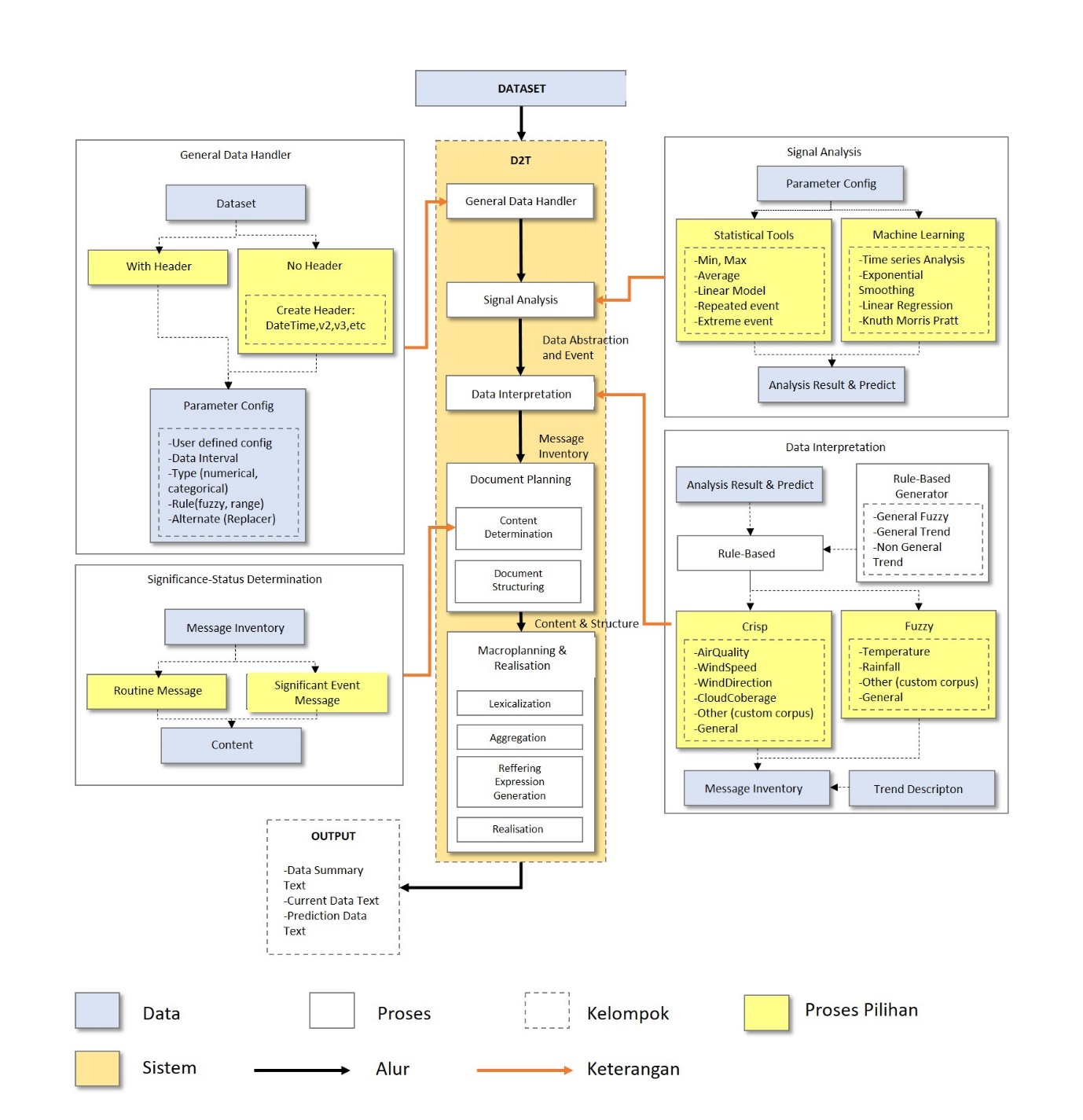
### Data Coba-coba

## Pengembangan Model *Data-to-*text untuk Kasus Data *General* dengan *Fuzzy Rule-based* dan *Machine Learning*

Model sistem *Data-to-Text* pada penelitian ini mengacu pada arsitektur sistem *Data-to-Text* yang dikembangkan oleh Reiter (2011), dan beberapa bagian yang dikembangkan oleh Putra (2017). Dengan menggunakan konsep *Fuzzy Rule-based* dan *Machine Learning* pada bagian *Signal Analysis* dan *Data Interpretation*, sistem yang dibangun dapat menerima masukan berupa data *general* atau data yang berasal dari bidang apapun baik data tersebut mempunyai informasi berupa *header*, tipe data, *rule* maupun tidak.

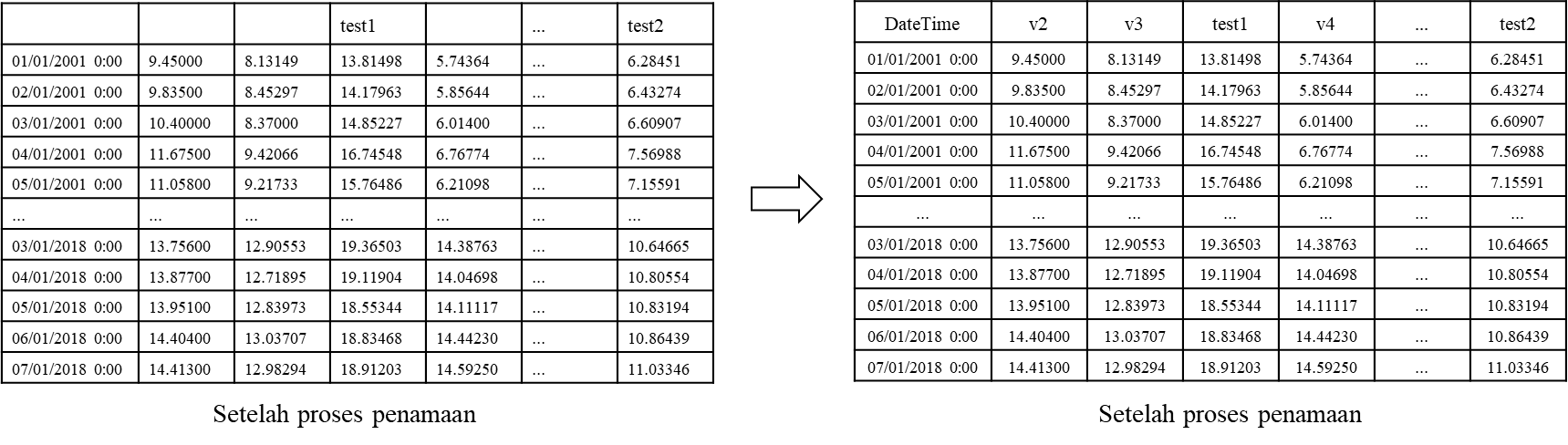
Keluaran yang dihasilkan oleh sistem ini berupa berita yang terdiri dari tiga bagian, bagian pertama merupakan deskripsi dari ringkasan data selama periode tertentu, bagian ke-dua merupakan deskripsi dari baris terakhir pada *n* buah data, bagian ke-tiga merupakan informasi dari hasil prediksi untuk baris data ke *n+1.*

Model penelitian pada Gambar 4.1 memperlihatkan perbedaan yang terdapat pada penelitian ini yaitu terdapatnya bagian *General Data Handler* dan pengaplikasian *Machine Learning* dan *Fuzzy Rule-based* pada bagian *Signal Analysis* dan *Data Interpretation*. Saat pertama kali data masuk, data akan diproses pada bagian *General Data Handler*, bagian ini akan mengelompokan apakah data memiliki informasi berupa *header* atau tidak, lalu sistem akan secara otomatis melihat apakah sudah terdapat konfigurasi untuk setiap parameter sesuai dengan *header* parameter tersebut, jika tidak terdapat konfigurasi maka sistem akan membuat konfigurasi dengan nilai *default*. Perbedaan lain terdapat pada bagian *Signal Analysis*, pada bagian ini digunakan *Machine Learning* dan *Time-Series Analysis* untuk mendapatkan pola diskret yang akan digunakan sebagai masukan untuk tahap *Data Interpretation*. Pola yang sudah didapatkan pada tahap *Signal Analysis* dimana pada proses *Data Interpretation* data akan diinterpretasikan sesuai tipe *Rule-based* yang terdapat pada konfigurasi parameter, dimana ada dua pilihan *Rule-based* untuk menginterpretasikan data yaitu, *Fuzzy* dan *Crisp.* Lalu dilanjutkan dengan serangkaian proses sehingga dihasilkan teks berita, serangkaian proses terebut dapat dilihat secara utuh pada kotak berwarna oranye.

Gambar 4. Model Data-to-text untuk Data *General*

## Model Proses *General Data Handler*

Proses *General Data Handler* bertujuan untuk pengolahan dan pemrosesan data, tahapan ini bertujuan agar dataset yang akan diproses menjadi *general* sebelum digunakan untuk proses selanjutnya. Pertama-tama data yang masuk akan dilihat terlebih dahulu, apakah dataset yang masuk setiap parameternya memiliki *header* atau tidak. Jika semua atau beberapa parameter pada suatu dataset ada yang tidak memiliki *header*, sistem akan memberi *header*  pada parameter tersebut secara otomatis. Khusus untuk *header pada* parameter atau kolom pertama dari dataset, sistem akan mengubah *header* menjadi DateTime. Untuk parameter selanjutnya yang tidak memiliki *header*, sistem akan menamai *header* dengan nama *v2, v3, v4,* dan seterusnya. Untuk lebih lengkapnya, bisa dilihat pada gambar 4.2 .



Gambar 4. Penamaan *header* pada Proses *General Data Handler*

Setelah penamaan *header*, proses selanjutnya yaitu konfigurasi parameter. Konfigurasi parameter merupakan aturan atau cara bagaimana suatu parameter akan diolah dan diproses dalam sistem. Konfigurasi parameter disimpan pada file *mainconfig.csv* yang terdapat pada folder *Config*. Terdapat 3 konfigurasi yang disimpan, yatu:

1. *Type,* konfigurasi ini akan menentukan bagaimana suatu parameter diproses. Terdapat dua kategori yaitu, *numerical*  dan *categorical*. Untuk parameter dengan tipe *numerical*, sistem akan melakukan proses analisis data menggunakan *Statistical Tools* seperti *min, max, average,* dan penerapan *Machine Learning* seperti *time-series analysis, regression, exponential smoothing,* dan lainnya*.* Untuk parameter dengan tipe *categorical,* sistem hanya akan melakukan pemrosesan *Repeated Event* dan penerapan *Motif Discovery* mengunakan algoritma *Knuth Morris Pratt* (KMP) pada parameter tersebut.
2. *Rule,* konfigurasi ini menentukan bagaimana nantinya suatu parameter akan diinterpretasikan. Untuk parameter dengan *rule fuzzy*, parameter tersebut akan diinterpretasikan menggunakan *Fuzzy membership function*. Sama halnya untuk parameter dengan *rule crisp,* maka parameter tersebutakan diinterpretasikan menggunakan *Crisp membership function.*
3. *Alternate*, konfigurasi ini dapat digunakan untuk mengubah nama parameter pada teks keluaran nantinya.

Tabel 4. Konfigurasi parameter pada file *mainconfig.csv*

|  |  |  |  |
| --- | --- | --- | --- |
| **ColName** | **Type** | **Rule** | **Alternate** |
| AirQuality | numeric | crisp | Air Quality |
| WindSpeed | numeric | crisp | Wind Speed |
| WindDirection | numeric | crisp | Wind Direction |
| CloudCoverage | numeric | crisp | Cloud Coverage |
| Temperature | numeric | fuzzy | NA |
| Rainfall | categorical | fuzzy | NA |
| USD | integer | NA | U.S. Dollar |
| JPY | NA | NA | Japan Yen |
| GBP | NA | NA | Great British Pounds |
| SGD | NA | NA | Singapore Dollar |
| MYR | NA | NA | Malaysian Ringgit |
| HKD | NA | NA | Hong Kong dollar |

Konfigurasi ini bersifat opsional, pengguna dapat mengubah atau menambahkan konfigurasi untuk suatu parameter pada file *mainconfig.csv,* lebih lengkapnya dapat dilihat pada tabel 4.4*.* Namun jika tidak ditemukan konfigurasi untuk suatu parameter pada dataset, maka sistem akan secara otomatis melakukan pemrosesan data secara *general* sesuai tipe datanya, lalu diinterpretasikan menggunakan *General Fuzzy membership function.* Tahapan lebih lengkapnya adalah sebagai berikut:

1. Buka file *mainconfig.csv* pada folder Config.
2. Pengguna dapat mengubah atau menambahkan konfigurasi dengan memasukan atau mengubah nama parameter, tipe data, *rule,* dan *alternate* pada file tersebut. Jika pengguna hanya mengisi beberapa konfigurasi saja, maka pengguna harus menuliskan NA untuk konfigurasi yang kosong atau tidak diisi.
3. Untuk konfigurasi *rule,* jika pengguna memasukan konfigurasi tersebut, maka pengguna harus menambahkan file pada folder Corpus dengan nama file sesuai dengan nama parameter yang diikuti dengan *Adjective.csv.* Seperti, *AirQualityAdjective.csv*, *TemperatureAdjective.csv,* dan *RainfallAdjective.csv.*

## Model Proses *Signal Analysis*

Proses *Signal Analysis* merupakan tahapan awal dari sistem D2T, tahapan ini bertujuan untuk mendeteksi dan menganalisis pola-pola diskret yang terdapat pada dataset. Selain itu, pada proses *Signal Analysis* ini dihasilkan juga ringkasan data seperti pada sistem *BABYTALK Family Sistem* yang dipaparkan oleh (Portet, dkk., 2009), sistem tersebut melakukan peringkasan data untuk mencari ringkasan kejadian selama 45 menit, dan juga pada sistem DWP yang menghasilkan ringkasan data klimatologi selama satu bulan dari data klimatologi selama satu tahun (Putra et al., 2017). Selain peringkasan data, pada tahap ini sistem juga melakukan analisis untuk mencari pola kejadian seperti *Repeated Event, Extreme Event,* analisis untuk data ke-*n* (data terakhir), dan prediksi data untuk data ke-*n+1*. Setelah melakukan serangkaian analisis pada tahap ini, hasil dari analasis tersebut akan disimpan dan digabungkan dalam bentuk *Data Abstraction and Event*. Untuk lebih jelasnya, input, proses, dan ouput pada tahap ini adalah sebagai berikut:

* Input : *Numerical Dataset with header*
* Proses : *Signal Analysis*
* Output : *Data Abstraction and Event*

Sub-proses pada tahap ini akan dibahas pada sub-bab berikutnya.

Putra, B., Riza, L. S., & Wihardi, Y. (2017). Pengembangan Sistem Data-to-Text untuk Membangkitkan Berita Cuaca dengan Pendekatan Time-Series dalam R.

Reiter, E. (2011). An Architecture for Data-to-Text Systems. *Computational Intelligence*, *27*(1), 23–40. https://doi.org/10.1111/j.1467-8640.2010.00370.x