



BAB III

ANALISIS SISTEM YANG BERJALAN

A. Gambaran Umum Objek Penelitian

PT Bank Rakyat Indonesia (BRI) adalah salah satu bank terbesar di Indonesia yang didirikan pada 16 Desember 1895. Bank ini berfokus pada pelayanan perbankan yang terjangkau dan inklusif bagi seluruh lapisan masyarakat di Indonesia. BRI memiliki peran penting dalam mendukung pengembangan sektor pertanian dan koperasi di Indonesia, serta memfasilitasi pertumbuhan ekonomi melalui pemberian kredit usaha kecil dan menengah (UKM) dan kredit perumahan.

BRI memiliki jaringan cabang yang luas di seluruh Indonesia dan memberikan kemudahan akses perbankan bagi masyarakat terutama yang berada di daerah terpencil dan pedesaan. Selain itu, BRI juga memiliki layanan perbankan elektronik yang dapat diakses melalui aplikasi mobile banking dan internet banking.

Kepercayaan dan kepuasan nasabah serta pemangku kepentingan adalah prioritas utama BRI. Bank ini berkomitmen untuk memberikan pelayanan terbaik dengan keamanan dan keandalan yang terjamin, serta mengikuti perkembangan teknologi terbaru dalam memberikan solusi finansial yang sesuai dengan kebutuhan nasabah.

B. Metodologi Penelitian

1. Teknik Pengumpulan Data

Dalam penelitian ini data yang digunakan adalah data sekunder dari Kaggle yaitu <https://www.kaggle.com/>. Alasan peneliti menggunakan data sekunder ini adalah untuk mempersingkat waktu pengumpulan data apabila pengumpulan data dilakukan secara manual akan memakan waktu cukup lama.



Dalam melakukan pengumpulan data, peneliti menggunakan metode kuantitatif, yaitu penelitian yang lebih sistematis, terencana, terstruktur, jelas dari awal hingga akhir penelitian serta tidak dipengaruhi oleh keadaan yang ada pada lapangan.

2. Teknik Analisis Data

Dalam penelitian ini, peneliti menggunakan teknik analisis yang dilakukan berdasarkan metode CRISP-DM (*Cross-Industry Standart Process for Data Mining*). CRISP-DM menyediakan proses standar yang tidak berpemilik dan tersedia secara bebas agar dapat menyesuaikan hasil *data mining* ke dalam strategi pemecahan masalah umum bisnis atau unit penelitian. CRISP-DM memiliki 6 siklus yang fasenya bersifat adaptif, antara lain:

- a. **Fase Bisnis/Pemahaman Penelitian (*Business/Research Understanding Phase*)**
 - (1) Pertama, menyampaikan dengan jelas tujuan dan persyaratan proyek dalam hal bisnis atau unit penelitian secara keseluruhan.
 - (2) Kemudian, terjemahkan tujuan dan batasan ini ke dalam perumusan definisi masalah *data mining*.
 - (3) Terakhir, siapkan strategi awal untuk mencapai tujuan yang telah ditetapkan.

Hak Cipta Dilindungi Undang-Undang

Institut Bisnis dan Informatika Kwik Kian Gie

1. Dilarang mengutip sebagian atau seluruh karya tulis ini tanpa mencantumkan dan menyebutkan sumber:
 - a. Pengutipan hanya untuk kepentingan pendidikan, penelitian, penulisan karya ilmiah, penyusunan laporan, penulisan kritik dan tinjauan suatu masalah.
 - b. Pengutipan tidak merugikan kepentingan yang wajar IBIKKG.
2. Dilarang mengumumkan dan memperbanyak sebagian atau seluruh karya tulis ini dalam bentuk apapun tanpa izin IBIKKG.



b. **Fase Pemahaman Data (*Data Understanding Phase*)**

- (1) Pertama, mengumpulkan data.
- (2) Kemudian, gunakan analisis data eksploratori untuk membiasakan diri dengan data, dan menemukan wawasan awal.
- (3) Mengevaluasi kualitas data.
- (4) Terakhir, memilih data yang sesuai dan berkemungkinan berisi pola agar dapat ditindaklanjuti.

c. **Fase Persiapan Data (*Data Preparation Phase*)**

- (1) Pada fase ini mencakup semua aspek mulai dari menyiapkan set data akhir, data awal, data mentah, dan data kotor.
- (2) Pilih kasus dan variabel yang diinginkan untuk melakukan proses analisis.
- (3) Melakukan transformasi pada variabel tertentu, jika diperlukan.
- (4) Bersihkan data mentah sehingga siap untuk alat pemodelan.

d. **Fase Model (*Modeling Phase*)**

- (1) Pilih dan terapkan teknik pemodelan yang sesuai.
- (2) Mengkalibrasi pengaturan model untuk mengoptimalkan hasil.
- (3) Sering kali, beberapa teknik yang berbeda dapat diterapkan untuk masalah penggalian data yang sama.
- (4) Mungkin memerlukan pengulangan kembali ke tahap persiapan data, untuk menyesuaikan bentuk data dengan persyaratan yang lebih spesifik dari teknik *data mining* tertentu.

© Hak cipta milik IBI KKG (Institut Bisnis dan Informatika Kwik Kian Gie)

Hak Cipta Dilindungi Undang-Undang

Institut Bisnis dan Informatika Kwik Kian Gie

1. Dilarang mengutip sebagian atau seluruh karya tulis ini tanpa mencantumkan dan menyebutkan sumber:
 - a. Pengutipan hanya untuk kepentingan pendidikan, penelitian, penulisan karya ilmiah, penyusunan laporan, penulisan kritik dan tinjauan suatu masalah.
 - b. Pengutipan tidak merugikan kepentingan yang wajar IBIKKG.
2. Dilarang mengumumkan dan memperbanyak sebagian atau seluruh karya tulis ini dalam bentuk apapun tanpa izin IBIKKG.



e. **Fase Evaluasi (*Evaluation Phase*)**

- (1) Fase pemodelan telah menghasilkan satu atau lebih model. Model-model ini harus dievaluasi kualitas dan efektivitasnya, sebelum digunakan di lapangan.
- (2) Menentukan apakah model tersebut sudah mencapai tujuan yang telah ditetapkan pada tahap pertama.
- (3) Menetapkan apakah beberapa aspek penting dari masalah bisnis atau penelitian belum diperhitungkan secara memadai.
- (4) Terakhir, buatlah keputusan mengenai penggunaan hasil *data mining*.

f. **Fase Penerapan (*Development Phase*)**

- (1) Pembuatan model tidak menandakan selesainya proyek. Perlu menggunakan model yang telah dibuat.
- (2) Contoh penerapan sederhana: Membuat laporan.
- (3) Contoh penerapan yang lebih kompleks: Menerapkan proses penggalian data paralel di departemen lain.
- (4) Untuk bisnis, pelanggan sering kali melakukan penerapan berdasarkan model yang telah digunakan

3. Penerapan Algoritma

Pada penelitian ini akan dibahas bagaimana penggunaan komparasi algoritma Long Short-Term Memory (LSTM) dan Regresi Linear untuk memprediksi harga saham BBRI dapat digambarkan pada diagram sebagai berikut:

© Hak cipta milik IBI KKG (Institut Bisnis dan Informatika Kwik Kian Gie)

Hak Cipta Dilindungi Undang-Undang

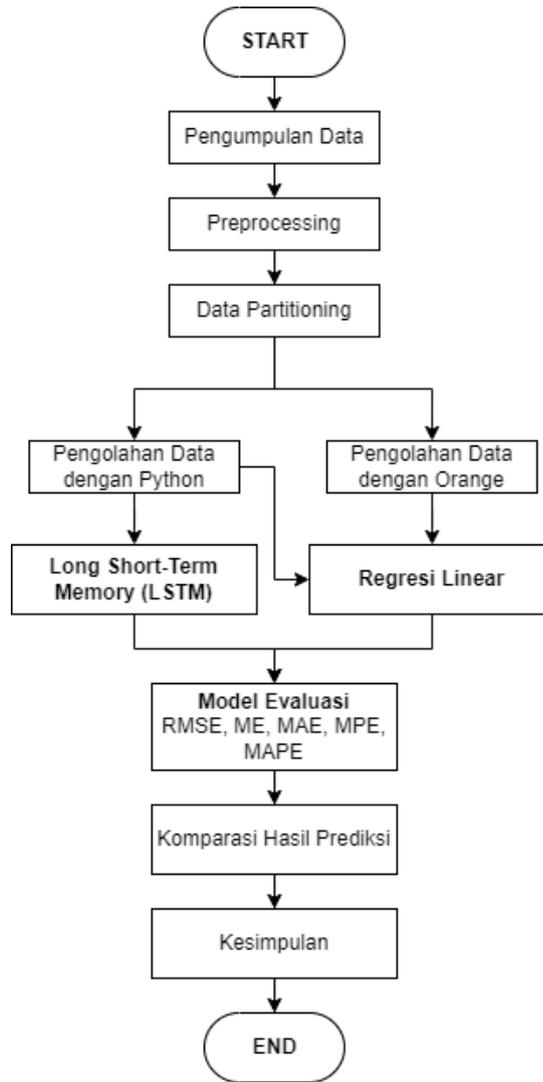
Institut Bisnis dan Informatika Kwik Kian Gie

1. Dilarang mengutip sebagian atau seluruh karya tulis ini tanpa mencantumkan dan menyebutkan sumber:
 - a. Pengutipan hanya untuk kepentingan pendidikan, penelitian, penulisan karya ilmiah, penyusunan laporan, penulisan kritik dan tinjauan suatu masalah.
 - b. Pengutipan tidak merugikan kepentingan yang wajar IBIKKG.
2. Dilarang mengumumkan dan memperbanyak sebagian atau seluruh karya tulis ini dalam bentuk apapun tanpa izin IBIKKG.



© Hak cipta milik IBI KKG (Institut Bisnis dan Informatika Kwik Kian Gie)

Hak Cipta Dilindungi Undang-Undang



Gambar 3.1
Workflow Penerapan Algoritma

Gambar 3.1 merupakan *workflow* yang digunakan oleh penulis untuk melakukan penelitian ini. Berikut adalah penjelasannya:

a. **Pengumpulan Data**

Pengumpulan data saham dikumpulkan secara sekunder melalui *website* <https://kaggle.com/>.

1. Dilarang mengutip sebagian atau seluruh karya tulis ini tanpa mencantumkan dan menyebutkan sumber:

a. Pengutipan hanya untuk kepentingan pendidikan, penelitian, penulisan karya ilmiah, penyusunan laporan, penulisan kritik dan tinjauan suatu masalah.

b. Pengutipan tidak merugikan kepentingan yang wajar IBIKKG.

2. Dilarang mengemukakan dan memperbanyak sebagian atau seluruh karya tulis ini dalam bentuk apapun tanpa izin IBIKKG.



© Hak cipta milik IBI KKG (Institut Bisnis dan Informatika Kwik Kian Gie)

Institut Bisnis dan Informatika Kwik Kian Gie

Hak Cipta Dilindungi Undang-Undang

- b. **Preprocessing**
Membersihkan data kotor menjadi data berkualitas untuk dilakukan tahap *data mining*.
- c. **Data Partitioning**
Melakukan pemisahan data antara *data train* dan *data test*.
- d. **Pengolahan Data dengan Python**
Pengolahan data menjadi hasil prediksi menggunakan Python.
- e. **Pengolahan Data dengan Orange**
Pengolahan data menjadi hasil prediksi menggunakan aplikasi Orange.
- f. **Long Short-Term Memory (LSTM)**
Data diolah menggunakan algoritma Long Short-Term Memory dengan bahasa pemrograman Python dan aplikasi Orange.
- g. **Regresi Linear**
Data diolah menggunakan algoritma Regresi Linear dengan aplikasi Orange.
- h. **Model Evaluasi**
Melakukan evaluasi dan mengkomparasi skor dari hasil pengolahan.
- i. **Komparasi Hasil Prediksi**
Melakukan komparasi keakuratan algoritma LSTM dan Regresi Linear menggunakan Python pada saham BBRI, dan melakukan komparasi terhadap algoritma Regresi Linear terhadap bahasa pemrograman Python dan aplikasi Orange.
- j. **Kesimpulan**
Kesimpulan dari hasil komparasi pada masing-masing algoritma.

1. Dilarang mengutip sebagian atau seluruh karya tulis ini tanpa mencantumkan dan menyebutkan sumber:

a. Pengutipan hanya untuk kepentingan pendidikan, penelitian, penulisan karya ilmiah, penyusunan laporan, penulisan kritik dan tinjauan suatu masalah.

b. Pengutipan tidak merugikan kepentingan yang wajar IBIKKG.

2. Dilarang menggunakan dan memperbanyak sebagian atau seluruh karya tulis ini dalam bentuk apapun tanpa izin IBIKKG.



4. Teknik Pengukuran Data

a. Pengukuran Kelayakan Metode Prediksi

Peneliti menggunakan Root Mean Square Error (RMSE) sebagai pengukuran kelayakan algoritma, untuk mengukur penyimpangan data prediksi, untuk menentukan tingkat kesalahan prediksi data berdasarkan pengaruh pada tiap variabel yang digunakan. Keakuratan pada pengukuran diestimasi dengan hasil RMSE memiliki nilai kecil yang mendekati nol. Pada gambar 3.2 merupakan gambaran rumus dasar RMSE.

$$RMSE = \sqrt{\frac{\sum_{t=1}^n (y_t - y)^2}{n}}$$

Gambar 3.2
Rumus Dasar RMSE

y_t = data prediksi

y = data actual

n = jumlah data

b. Pengukuran Penyimpangan Prediksi

Peneliti menggunakan Mean Error (ME), Mean Absolute Error (MAE), Mean Percentage Error (MPE), dan Mean Absolute Percentage Error (MAPE) untuk mengukur penyimpangan data prediksi. Berikut adalah penjelasan dari masing-masing metode pengukuran penyimpangan prediksi yang peneliti gunakan:

© Hak cipta milik IBI KKG (Institut Bisnis dan Informatika Kwik Kian Gie)

Hak Cipta Dilindungi Undang-Undang

Institut Bisnis dan Informatika Kwik Kian Gie

1. Dilarang mengutip sebagian atau seluruh karya tulis ini tanpa mencantumkan dan menyebutkan sumber:
 - a. Pengutipan hanya untuk kepentingan pendidikan, penelitian, penulisan karya ilmiah, penyusunan laporan, penulisan kritik dan tinjauan suatu masalah.
 - b. Pengutipan tidak merugikan kepentingan yang wajar IBIKKG.
2. Dilarang mengumumkan dan memperbanyak sebagian atau seluruh karya tulis ini dalam bentuk apapun tanpa izin IBIKKG.



Hak Cipta Dilindungi Undang-Undang

1. Dilarang mengutip sebagian atau seluruh karya tulis ini tanpa mencantumkan dan menyebutkan sumber:
 - a. Pengutipan hanya untuk kepentingan pendidikan, penelitian, penulisan karya ilmiah, penyusunan laporan, penulisan kritik dan tinjauan suatu masalah.
 - b. Pengutipan tidak merugikan kepentingan yang wajar IBIKKG.
2. Dilarang mengumumkan dan memperbanyak sebagian atau seluruh karya tulis ini dalam bentuk apapun tanpa izin IBIKKG.

(1) *Mean Error (ME)*

Rata-rata dari margin selisih antara data yang diprediksi dengan data aktual yang menghasilkan seberapa besar nilai kesalahan prediksi. Rumus perhitungan Mean Error dapat dilihat pada gambar 3.3.

$$ME = \frac{\sum E}{n}$$

Gambar 3.3
Rumus Dasar Mean Error

E = Data Aktual – Data Prediksi

n = Jumlah Data

(2) *Mean Absolute Error (MAE)*

Rata-rata absolut dari margin selisih antara data yang diprediksi dengan data aktual yang menghasilkan seberapa besar nilai kesalahan prediksi. Rumus perhitungan Mean Absolute Error dapat dilihat pada gambar 3.4.

$$MAE = \left| \frac{\sum E}{n} \right|$$

Gambar 3.4
Rumus Dasar Mean Absolute Error

E = Data Aktual – Data Prediksi

n = Jumlah Data



c. Algoritma Long Short-Term Memory

Alih-alih unit yang hanya menerapkan nonlinieritas elementer pada transformasi affine dari input dan unit berulang, jaringan berulang LSTM memiliki "sel LSTM" yang memiliki perulangan internal (loop mandiri), di samping perulangan eksternal RNN. Setiap sel memiliki input dan output yang sama dengan jaringan berulang biasa, tetapi memiliki lebih banyak parameter dan sistem *unit gating* yang mengontrol aliran informasi. Gambar 3.5 merupakan formula dari LSTM:

$$f_i^{(t)} = \sigma \left(b_i^f + \sum_j U_{i,j}^f x_j^{(t)} + \sum_j W_{i,j}^f h_j^{(t-1)} \right),$$

Gambar 3.5

Rumus Forget Gate

Ian Goodfellow, Yoshua Bengio, et al (2018:410)

$x^{(t)}$ adalah vektor input saat ini dan $h^{(t)}$ adalah vektor lapisan tersembunyi saat ini, yang berisi output dari semua sel LSTM, dan b^f, U^f, W^f adalah bias, bobot masukan, dan bobot berulang untuk *forget gates*. Status internal sel LSTM dapat diperbarui, tetapi dengan bobot loop mandiri bersyarat $f_i^{(t)}$. Gambar 3.6 merupakan rumus *self-loop weight*:

$$s_i^{(t)} = f_i^{(t)} s_i^{(t-1)} + g_i^{(t)} \sigma \left(b_i + \sum_j U_{i,j} x_j^{(t)} + \sum_j W_{i,j} h_j^{(t-1)} \right),$$

Gambar 3.6

Rumus Self-loop Weight

Ian Goodfellow, Yoshua Bengio, et al (2018:411)

b, U dan W masing-masing menunjukkan bias, bobot masukan, dan bobot berulang ke dalam sel LSTM. Unit gerbang input eksternal $g_i^{(t)}$ dihitung

1. Dilarang mengutip sebagian atau seluruh karya tulis ini tanpa mencantumkan dan menyebutkan sumber:
 - a. Pengutipan hanya untuk kepentingan pendidikan, penelitian, penulisan karya ilmiah, penyusunan laporan, penulisan kritik dan tinjauan suatu masalah.
 - b. Pengutipan tidak merugikan kepentingan yang wajar IBIKKG.
2. Dilarang mengumumkan dan memperbanyak sebagian atau seluruh karya tulis ini dalam bentuk apapun tanpa izin IBIKKG.



1. Dilarang mengutip sebagian atau seluruh karya tulis ini tanpa mencantumkan dan menyebutkan sumber:
 - a. Pengutipan hanya untuk kepentingan pendidikan, penelitian, penulisan karya ilmiah, penyusunan laporan, penulisan kritik dan tinjauan suatu masalah.
 - b. Pengutipan tidak merugikan kepentingan yang wajar IBIKKG.
2. Dilarang menggunakan dan memperbanyak sebagian atau seluruh karya tulis ini dalam bentuk apapun tanpa izin IBIKKG.

dengan cara yang sama dengan *forget gates* (dengan unit sigmoid untuk mendapatkan nilai gerbang antara 0 dan 1), tetapi dengan parameternya sendiri. Gambar 3.7 merupakan formula untuk *external input gate*:

$$g_i^{(t)} = \sigma \left(b_i^g + \sum_j U_{i,j}^g x_j^{(t)} + \sum_j W_{i,j}^g h_j^{(t-1)} \right).$$

Gambar 3.7
Rumus External Input Gate
Ian Goodfellow, Yoshua Bengio, et al (2018:411)

Keluaran $h_i^{(t)}$ dari sel LSTM juga dapat dimatikan, melalui gerbang keluaran $q_i^{(t)}$, yang juga menggunakan unit sigmoid untuk gating. Gambar 3.8 merupakan formula rumus *output gate*:

$$h_i^{(t)} = \tanh \left(s_i^{(t)} \right) q_i^{(t)}$$

$$q_i^{(t)} = \sigma \left(b_i^o + \sum_j U_{i,j}^o x_j^{(t)} + \sum_j W_{i,j}^o h_j^{(t-1)} \right)$$

Gambar 3.8
Rumus Output Gate
Ian Goodfellow, Yoshua Bengio, et al (2018:411)

Parameter b^o, U^o, W^o memiliki bias, bobot input, dan bobot berulang. Di antara variannya, dapat memilih untuk menggunakan status sel $i - th$. Namun untuk melakukan hal ini 3 parameter tambahan diperlukan.

Jaringan LSTM dapat mempelajari ketergantungan jangka panjang dengan lebih mudah dibandingkan dengan arsitektur berulang yang sederhana, pertama-tama pada kumpulan data buatan yang dirancang untuk menguji kemampuan mempelajari ketergantungan jangka panjang.



d. Algoritma Regresi Linear

Regresi Linear Sederhana adalah sebuah metode sederhana untuk memprediksi nilai variabel dependen Y berdasarkan satu variabel independen X. Berikut adalah penulisan formula regresi linier sederhana:

$$Y \approx b_0 + b_1X$$

Y : atau dependent variable, adalah akibat dari suatu sebab, misalnya: kenaikan gaji karyawan (terhadap lama pengalaman kerja), atau tinggi/rendah ranking murid (terhadap lama waktu belajarnya).

X : atau independent variable, adalah hal yang diasumsikan menjadi sebab atas suatu hal, yang nantinya bisa mengakibatkan Y .

b_1 : atau coefficient, adalah suatu unit / proporsi yang dapat mengubah nilai “ x ”

b_0 : atau constant, adalah nilai awal “ x ” pada suatu kejadian.

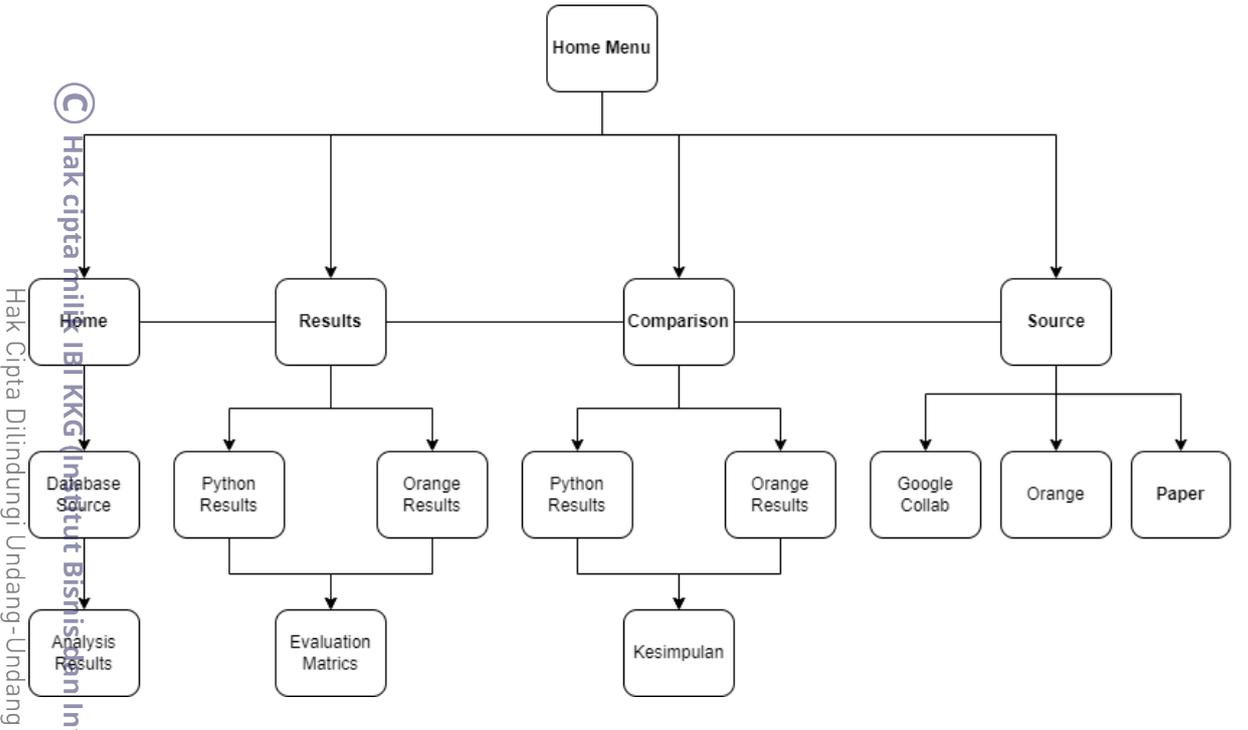
C Hak cipta milik IBI KKG (Institut Bisnis dan Informatika Kwik Kian Gie)

Hak Cipta Dilindungi Undang-Undang

C. Teknik Perancangan Graphic User Interface (GUI)

Pada perancangan Graphic User Interface (GUI), peneliti menggunakan Google Sites untuk menyajikan hasil proses data mining dalam bentuk dashboard. Pada halaman website terdapat Page Home, Page Results, Page Comparison, dan Page Source. Berikut gambar hirarki dari rancangan GUI tersebut :

1. Dilarang mengutip sebagian atau seluruh karya tulis ini tanpa mencantumkan dan menyebutkan sumber:
 - a. Pengutipan hanya untuk kepentingan pendidikan, penelitian, penulisan karya ilmiah, penyusunan laporan, penulisan kritik dan tinjauan suatu masalah.
 - b. Pengutipan tidak merugikan kepentingan yang wajar IBIKKG.
2. Dilarang mengemukakan dan memperbanyak sebagian atau seluruh karya tulis ini dalam bentuk apapun tanpa izin IBIKKG.



Gambar 3.9
Hierarki Rancangan GUI

© Hak cipta milik IBIKKG (Institut Bisnis dan Informatika Kwik Kian Gie)
Hak Cipta Dilindungi Undang-Undang

Institut Bisnis dan Informatika Kwik Kian Gie

1. Dilarang mengutip sebagian atau seluruh karya tulis ini tanpa mencantumkan dan menyebutkan sumber:
 - a. Pengutipan hanya untuk kepentingan pendidikan, penelitian, penulisan karya ilmiah, penyusunan laporan, penulisan kritik dan tinjauan suatu masalah.
 - b. Pengutipan tidak merugikan kepentingan yang wajar IBIKKG.
2. Dilarang mengumumkan dan memperbanyak sebagian atau seluruh karya tulis ini dalam bentuk apapun tanpa izin IBIKKG.