

HOME  
([HTTPS://LIBRARY.GUNADARMA.AC.ID/DEPC\\_SYSTEM/](https://library.gunadarma.ac.id/DEPC_SYSTEM/))

PEMBAYARAN  
([HTTPS://LIBRARY.GUNADARMA.AC.ID/DEPC\\_SYSTEM/PEMBAYARAN](https://library.gunadarma.ac.id/DEPC_SYSTEM/PEMBAYARAN))

BIODATA DIRI  
([HTTPS://LIBRARY.GUNADARMA.AC.ID/DEPC\\_SYSTEM/ANGGOTA](https://library.gunadarma.ac.id/DEPC_SYSTEM/ANGGOTA))

PEMINJAMAN BUKU  
([HTTPS://LIBRARY.GUNADARMA.AC.ID/DEPC\\_SYSTEM/ANGGOTA/PEMINJAMAN](https://library.gunadarma.ac.id/DEPC_SYSTEM/ANGGOTA/PEMINJAMAN))

PENULISAN ILMIAH  
([HTTPS://LIBRARY.GUNADARMA.AC.ID/DEPC\\_SYSTEM/ANGGOTA/PENULISAN](https://library.gunadarma.ac.id/DEPC_SYSTEM/ANGGOTA/PENULISAN))

FOTOCOPY  
([HTTPS://LIBRARY.GUNADARMA.AC.ID/DEPC\\_SYSTEM/ANGGOTA/FOTOCOPY](https://library.gunadarma.ac.id/DEPC_SYSTEM/ANGGOTA/FOTOCOPY))

SUMBANG BUKU  
([HTTPS://LIBRARY.GUNADARMA.AC.ID/DEPC\\_SYSTEM/ANGGOTA/SUMBANGBUKU](https://library.gunadarma.ac.id/DEPC_SYSTEM/ANGGOTA/SUMBANGBUKU))

KASUS & NOTIFIKASI  
([HTTPS://LIBRARY.GUNADARMA.AC.ID/DEPC\\_SYSTEM/ANGGOTA/NOTIFIKASI](https://library.gunadarma.ac.id/DEPC_SYSTEM/ANGGOTA/NOTIFIKASI))

BEBAS PERPUSTAKAAN  
([HTTPS://LIBRARY.GUNADARMA.AC.ID/DEPC\\_SYSTEM/ANGGOTA/BEBASPERPUSTAKAAN](https://library.gunadarma.ac.id/DEPC_SYSTEM/ANGGOTA/BEBASPERPUSTAKAAN))

KUNJUNGAN  
([HTTPS://LIBRARY.GUNADARMA.AC.ID/DEPC\\_SYSTEM/KUNJUNGAN](https://library.gunadarma.ac.id/DEPC_SYSTEM/KUNJUNGAN))

E-PAPER  
([HTTPS://LIBRARY.GUNADARMA.AC.ID/DEPC](https://library.gunadarma.ac.id/DEPC))

FILKOM/KA/SKRIPSI/5209/2022

## IMPLEMENTASI MACHINE LEARNING MENGGUNAKAN ALGORITMA RANDOM FOREST REGRESSION UNTUK PREDIKSI HARGA JUAL RUMAH

EGA SRI LESTARI | 12118163 | 2022 | S1 - SKRIPSI | Sistem Informasi - Ilmu Komputer dan Teknologi Informasi

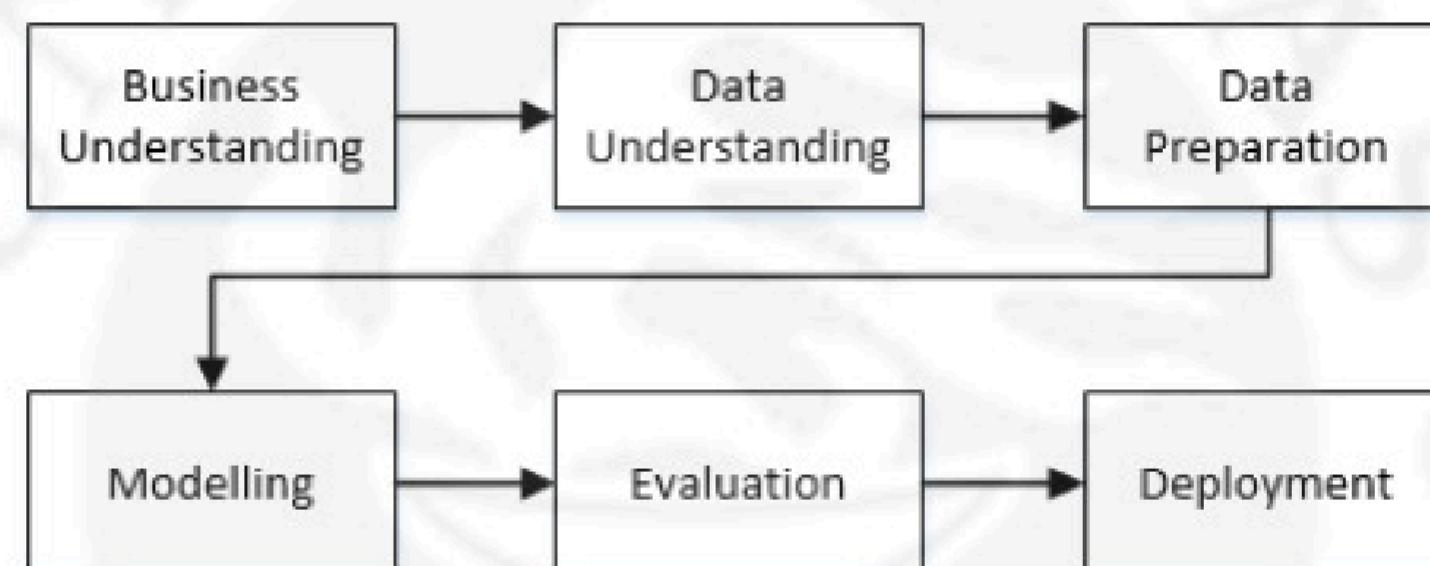
Pembimbing: Ida Astuti | Promotor:

### Table Of Contents

ABSTRAKSI (BAHASA INGGRIS)	KATA PENGANTAR	DAFTAR ISI	DAFTAR TABEL	DAFTAR GAMBAR	DAFTAR LAMPIRAN	BAB I	BAB II	<b>BAB III</b>	BAB IV	BAB V	DAFTAR P
----------------------------	----------------	------------	--------------	---------------	-----------------	-------	--------	----------------	--------	-------	----------

### 3. METODE PENELITIAN

Pada bab metode penelitian berisi tahapan perencanaan dan alur dari penelitian yang akan dilakukan. Metode penelitian ini akan menggunakan salah satu metodologi *data science* yaitu CRISP-DM (*Cross-Industry Standard Process for Data Mining*). Aplikasi yang akan dibangun pada penelitian ini berupa implementasi *Machine Learning* berbasis *website*. Proses yang akan dilakukan penelitian ini dapat dilihat pada Gambar 3.1.



Gambar 3. 1 Tahapan Penelitian

Berdasarkan Gambar 3.1 di atas merupakan proses tahapan penelitian akan dijabarkan sebagai berikut:

1. Pada tahap pertama adalah *business understanding* dimana akan melakukan pengidentifikasi terhadap masalah yang akan di selesaikan dan menentukan model *analytics* termasuk untuk menghitung *performance* dari model tersebut.
2. Proses selanjutnya, data *understanding* akan melakukan proses pengambilan dan pengumpulan data. Teknik pengambilan data akan dilakukan menggunakan *web scraping* pada *website* rumah123.com dengan *library BeautifulSoup Python* kemudian data akan disimpan dalam bentuk file berekstensi csv. Pada proses ini juga melakukan validasi data dengan

- [!\[\]\(95b42f0077faf7439a26242a54e021ec\_img.jpg\) HOME  
\(HTTPS://LIBRARY.GUNADARMA.AC.ID/DEPC  
SYSTEM/\)](https://library.gunadarma.ac.id/depc/home)
- [!\[\]\(e097ab4c08b8186dd0908330bbc2dc28\_img.jpg\) PEMBAYARAN  
\(HTTPS://LIBRARY.GUNADARMA.AC.ID/DEPC  
SYSTEM/PEMBAYARAN\)](https://library.gunadarma.ac.id/depc/pembayaran)
- [!\[\]\(1e9d865c5de095f8e3304757c49e79d7\_img.jpg\) BIODATA DIRI  
\(HTTPS://LIBRARY.GUNADARMA.AC.ID/DEPC  
SYSTEM/ANGGOTA\)](https://library.gunadarma.ac.id/depc/biodata)
- [!\[\]\(735b10d724a5f0ec5005c4eb3eb9c9d1\_img.jpg\) PEMINJAMAN BUKU  
\(HTTPS://LIBRARY.GUNADARMA.AC.ID/DEPC  
SYSTEM/ANGGOTA/PEMINJAMAN\)](https://library.gunadarma.ac.id/depc/peminjaman)
- [!\[\]\(e6250f05bc27fa93236b816562b699f9\_img.jpg\) PENULISAN ILMIAH  
\(HTTPS://LIBRARY.GUNADARMA.AC.ID/DEPC  
SYSTEM/ANGGOTA/PENULISAN\)](https://library.gunadarma.ac.id/depc/penulisan)
- [!\[\]\(d190cc638f389909d4b049d6c19e4cb2\_img.jpg\) FOTOCOPY  
\(HTTPS://LIBRARY.GUNADARMA.AC.ID/DEPC  
SYSTEM/ANGGOTA/FOTOCOPY\)](https://library.gunadarma.ac.id/depc/fotocopy)
- [!\[\]\(4d34001966c7597a8c3e4293694bde37\_img.jpg\) SUMBANG BUKU  
\(HTTPS://LIBRARY.GUNADARMA.AC.ID/DEPC  
SYSTEM/ANGGOTA/SUMBANGBUKU\)](https://library.gunadarma.ac.id/depc/sumbang)
- [!\[\]\(70b9adae95aa76ce55f26a9fb944efce\_img.jpg\) KASUS & NOTIFIKASI  
\(HTTPS://LIBRARY.GUNADARMA.AC.ID/DEPC  
SYSTEM/ANGGOTA/NOTIFIKASI\)](https://library.gunadarma.ac.id/depc/kasus)
- [!\[\]\(7b780ea9fe954e20f3a1890df3f944e2\_img.jpg\) BEBAS PERPUSTAKAAN  
\(HTTPS://LIBRARY.GUNADARMA.AC.ID/DEPC  
SYSTEM/ANGGOTA/BEBASPERPUSTAKAAN\)](https://library.gunadarma.ac.id/depc/bebas)
- [!\[\]\(6ed942f5426e79452fae08b09c55d70a\_img.jpg\) KUNJUNGAN  
\(HTTPS://LIBRARY.GUNADARMA.AC.ID/DEPC  
SYSTEM/KUNJUNGAN\)](https://library.gunadarma.ac.id/depc/kunjungan)
- [!\[\]\(7ac021fc5aa83a899c22bb9c073b3a95\_img.jpg\) E-PAPER  
\(HTTPS://LIBRARY.GUNADARMA.AC.ID/DEPC](https://library.gunadarma.ac.id/depc/e-paper)

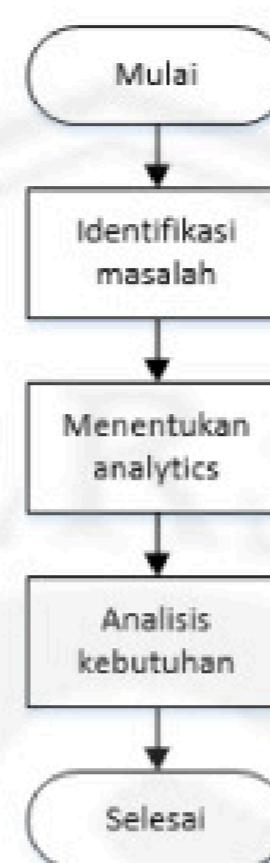
menilai kesesuaian kualitas data, menganalisa data secara eksploratif dengan visualisasi.

3. Tahap ketiga adalah data *preparation* yang berfungsi untuk memperbaiki kualitas pada data untuk melanjutkan pada proses pemodelan. Melakukan pemilihan data yang akan digunakan, membersihkan data yang tidak lengkap atau salah, dan melakukan konstruksi data dengan mentransformasikan data dan normalisasi data.
4. Tahap keempat adalah *modeling* yang akan mengembangkan model dengan pola data yang sudah dilakukan tahap konstruksi data. Memodelkan data *training* dengan menggunakan algoritma *Random Forest Regression* jika sudah selesai melakukan pelatihan maka akan dilanjutkan dengan proses pelatihan menggunakan algoritma *Cosine Similarity* untuk memperoleh hasil rekomendasi. *Tuning parameter* digunakan untuk meningkatkan hasil model *random forest* dengan parameter terbaik.
5. Pada tahap evaluasi akan melakukan akurasi data pada hasil prediksi dan rekomendasi yang diberikan menggunakan data *testing* menggunakan beberapa metrik seperti MSE, MAE dan *R-Squared*.
6. Tahap *deployment* dimana akan mengimplementasikan hasil model yang telah disimpan dan hasil rekomendasi ke dalam *website* menggunakan *framework* yang disediakan *Python* yaitu *flask*. Tahap terakhir pada penelitian adalah pengujian *website* prediksi harga rumah yang akan menggunakan pengujian inferensial dan pengujian *user acceptance*.

### 3.1. Business Understanding

Implementasi *Machine Learning* pada penelitian ini akan melakukan pemahaman kebutuhan serta tujuan untuk pengembangan bisnis selanjutnya. Masalah yang akan diselesaikan adalah mewujudkan hasil prediksi rumah berdasarkan inputan pengguna serta memberikan rekomendasi yang berkaitan dengan inputan tersebut.

- [HOME  
\(HTTPS://LIBRARY.GUNADARMA.AC.ID/DEPC SYSTEM/\)](#)
- [PEMBAYARAN  
\(HTTPS://LIBRARY.GUNADARMA.AC.ID/DEPC SYSTEM/PEMBAYARAN\)](#)
- [BIODATA DIRI  
\(HTTPS://LIBRARY.GUNADARMA.AC.ID/DEPC SYSTEM/ANGGOTA\)](#)
- [PEMINJAMAN BUKU  
\(HTTPS://LIBRARY.GUNADARMA.AC.ID/DEPC SYSTEM/ANGGOTA/PEMINJAMAN\)](#)
- [PENULISAN ILMIAH  
\(HTTPS://LIBRARY.GUNADARMA.AC.ID/DEPC SYSTEM/ANGGOTA/PENULISAN\)](#)
- [FOTOCOPY  
\(HTTPS://LIBRARY.GUNADARMA.AC.ID/DEPC SYSTEM/ANGGOTA/FOTOCOPY\)](#)
- [SUMBANG BUKU  
\(HTTPS://LIBRARY.GUNADARMA.AC.ID/DEPC SYSTEM/ANGGOTA/SUMBANGBUKU\)](#)
- [KASUS & NOTIFIKASI  
\(HTTPS://LIBRARY.GUNADARMA.AC.ID/DEPC SYSTEM/ANGGOTA/NOTIFIKASI\)](#)
- [BEBAS PERPUSTAKAAN  
\(HTTPS://LIBRARY.GUNADARMA.AC.ID/DEPC SYSTEM/ANGGOTA/BEBAisperpusTAKAAN\)](#)
- [KUNJUNGAN  
\(HTTPS://LIBRARY.GUNADARMA.AC.ID/DEPC SYSTEM/KUNJUNGAN\)](#)
- [E-PAPER  
\(HTTPS://LIBRARY.GUNADARMA.AC.ID/DEPC](#)

Gambar 3. 2 Proses *Business Understanding*

Berdasarkan Gambar 3.2 penelitian ini akan mengembangkan sebuah inovasi bisnis dalam proses memprediksi harga rumah dengan aplikasi berbasis *website* menggunakan *Machine Learning* dan memberikan hasil rekomendasi dari beberapa rumah yang memiliki kemiripan terhadap inputan pengguna. Metode yang cocok digunakan untuk menyelesaikan masalah ini adalah menggunakan algoritma regresi yang dapat memprediksi nilai terhadap inputan yang diberikan serta menggunakan algoritma perhitungan jarak agar menghasilkan rekomendasi yang dapat disesuaikan dengan kemiripan data input. Melakukan sebuah penelitian prediksi harga rumah berdasarkan beberapa kriteria ada beberapa kebutuhan yang diperlukan.

Tabel 3. 1 Kebutuhan Data

No.	Atribut	Keterangan
1.	Nama	Nama rumah yang dijual
2.	Harga	Harga rumah yang dijual
3.	Kota	Kota rumah yang dijual
4.	Kamar tidur	Jumlah kamar tidur pada rumah yang dijual
5.	Kamar mandi	Jumlah kamar mandi pada rumah yang dijual

HOME  
(HTTPS://LIBRARY.GUNADARMA.AC.ID/DEPC  
SYSTEM/)

PEMBAYARAN  
(HTTPS://LIBRARY.GUNADARMA.AC.ID/DEPC  
SYSTEM/PEMBAYARAN)

BIODATA DIRI  
(HTTPS://LIBRARY.GUNADARMA.AC.ID/DEPC  
SYSTEM/ANGGOTA)

PEMINJAMAN BUKU  
(HTTPS://LIBRARY.GUNADARMA.AC.ID/DEPC  
SYSTEM/ANGGOTA/PEMINJAMAN)

PENULISAN ILMIAH  
(HTTPS://LIBRARY.GUNADARMA.AC.ID/DEPC  
SYSTEM/ANGGOTA/PENULISAN)

FOTOCOPY  
(HTTPS://LIBRARY.GUNADARMA.AC.ID/DEPC  
SYSTEM/ANGGOTA/FOTOCOPY)

SUMBANG BUKU  
(HTTPS://LIBRARY.GUNADARMA.AC.ID/DEPC  
SYSTEM/ANGGOTA/SUMBANGBUKU)

KASUS & NOTIFIKASI  
(HTTPS://LIBRARY.GUNADARMA.AC.ID/DEPC  
SYSTEM/ANGGOTA/NOTIFIKASI)

BEBAS PERPUSTAKAAN  
(HTTPS://LIBRARY.GUNADARMA.AC.ID/DEPC  
SYSTEM/ANGGOTA/BEBASTERPUSTAKAAN)

KUNJUNGAN  
(HTTPS://LIBRARY.GUNADARMA.AC.ID/DEPC  
SYSTEM/KUNJUNGAN)

E-PAPER  
(HTTPS://LIBRARY.GUNADARMA.AC.ID/DEPC

6.	Garasi	Jumlah muatan mobil dalam garasi rumah
7.	Luas tanah	Luas tanah pada rumah yang dijual
8.	Luas bangunan	Luas bangunan pada rumah yang dijual
9.	Link	Link informasi lengkap mengenai rumah yang dijual

Berdasarkan Tabel 3.1 di atas merupakan kriteria dari atribut yang harus dimiliki pada saat pengumpulan data. Data yang akan dibutuhkan dalam menyelesaikan masalah ini diantaranya data nama rumah, 1 data target yaitu harga, 6 data *predictor* yaitu kota, kamar mandi, kamar tidur, garasi, luas tanah, dan luas bangunan, sedangkan data link digunakan untuk informasi data rekomendasi.

Tabel 3. 2 Kebutuhan Perangkat Keras

No.	Komponen	Spesifikasi
1.	Processor	Intel® Core™ i3-7020U CPU @ 2.30GHz (4CPUs), ~2.3GHz
2.	RAM	4GB
3.	OS Type	Windows 10 Home Single Language 64-bit
4.	Name	ASUS Vivobook 14 X407UAR

Berdasarkan Tabel 3.2 di atas merupakan kebutuhan perangkat keras pada penelitian ini. Perangkat keras yang digunakan akan menunjang keberhasilan yang dilakukan pada penelitian, jadi data tersebut merupakan batas minimal perangkat keras yang digunakan untuk proses penelitian.

Tabel 3. 3 Kebutuhan Perangkat Lunak

No.	Perangkat Lunak	Fungsi
1.	Google Colaboratory	Pengolahan big data dengan <i>Machine Learning</i>
2.	Microsoft Excel 2016	Pengolahan data sampel perhitungan manual
3.	Microsoft Visio 2013	Pembuatan flowchart dan struktur navigasi

HOME  
(HTTPS://LIBRARY.GUNADARMA.AC.ID/DEPC  
SYSTEM/)

PEMBAYARAN  
(HTTPS://LIBRARY.GUNADARMA.AC.ID/DEPC  
SYSTEM/PEMBAYARAN)

BIODATA DIRI  
(HTTPS://LIBRARY.GUNADARMA.AC.ID/DEPC  
SYSTEM/ANGGOTA)

PEMINJAMAN BUKU  
(HTTPS://LIBRARY.GUNADARMA.AC.ID/DEPC  
SYSTEM/ANGGOTA/PEMINJAMAN)

PENULISAN ILMIAH  
(HTTPS://LIBRARY.GUNADARMA.AC.ID/DEPC  
SYSTEM/ANGGOTA/PENULISAN)

FOTOCOPY  
(HTTPS://LIBRARY.GUNADARMA.AC.ID/DEPC  
SYSTEM/ANGGOTA/FOTOCOPY)

SUMBANG BUKU  
(HTTPS://LIBRARY.GUNADARMA.AC.ID/DEPC  
SYSTEM/ANGGOTA/SUMBANGBUKU)

KASUS & NOTIFIKASI  
(HTTPS://LIBRARY.GUNADARMA.AC.ID/DEPC  
SYSTEM/ANGGOTA/NOTIFIKASI)

BEBAS PERPUSTAKAAN  
(HTTPS://LIBRARY.GUNADARMA.AC.ID/DEPC  
SYSTEM/ANGGOTA/BEBAZPERPUSTAKAAN)

KUNJUNGAN  
(HTTPS://LIBRARY.GUNADARMA.AC.ID/DEPC  
SYSTEM/KUNJUNGAN)

E-PAPER  
(HTTPS://LIBRARY.GUNADARMA.AC.ID/DEPC

4.	Microsoft Power BI	Pembuatan visualisasi data
5.	Visual Studio Code	Pengembangan sistem
6.	Anaconda3 dan Python 3.9	Bahasa pemrograman
7.	Microsoft Word	Pengolahan laporan penelitian
8.	PyCharm	Pengambilan data web scraping

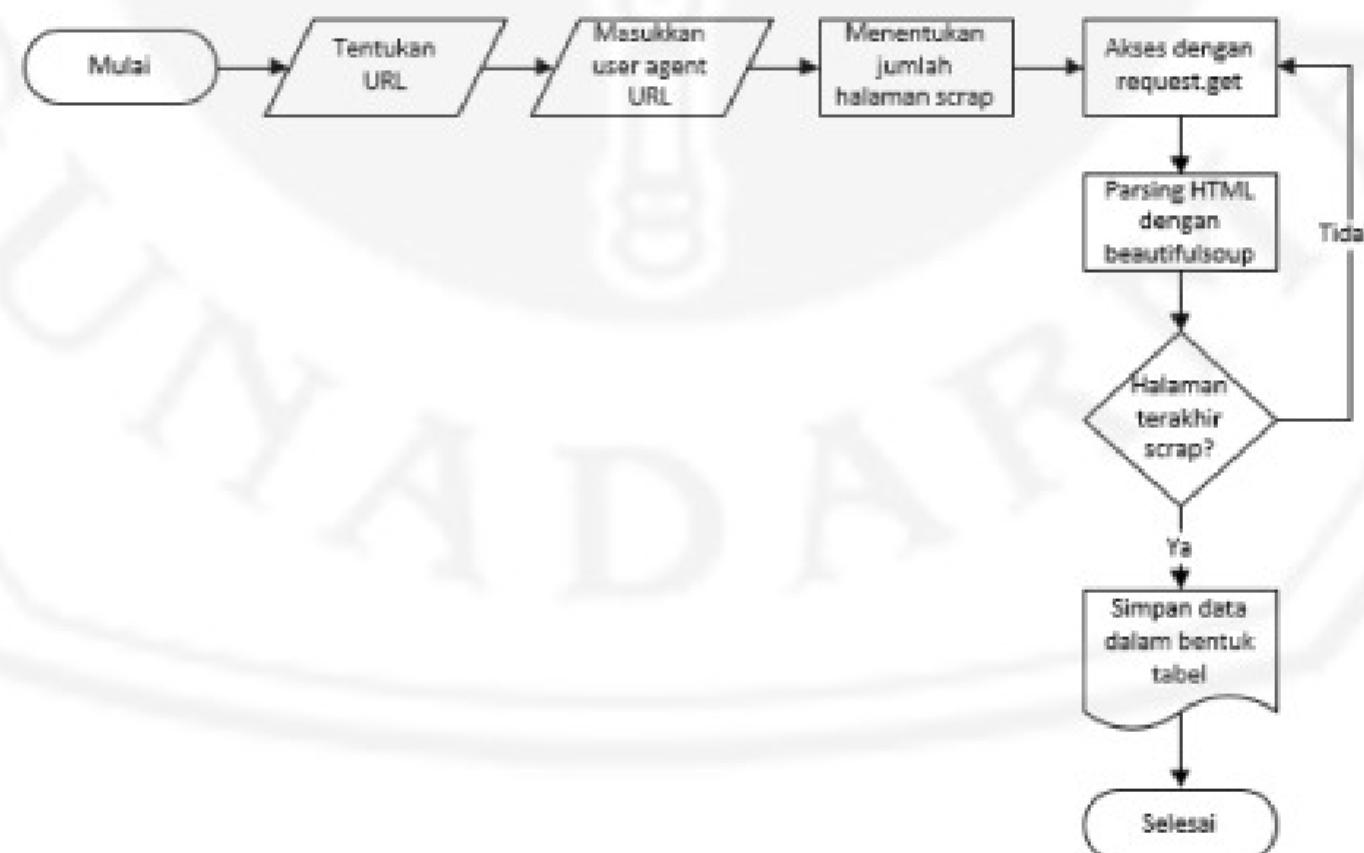
Berdasarkan Tabel 3.3 merupakan kebutuhan perangkat lunak yang akan digunakan pada penelitian ini. Ada begitu banyak perangkat lunak yang dapat digunakan selain 8 perangkat lunak di atas, jadi untuk penggunaan perangkat lunak dapat disesuaikan dengan kebutuhan masing-masing.

### 3.2. Data *Understanding*

Pada tahap data *understanding* penelitian ini akan dilakukan beberapa proses seperti pengumpulan data, data *validation*, menelaah data, dan eksplorasi data.

#### 3.2.1. Pengumpulan Data

Proses pengambilan data ini menggunakan sumber data eksternal yaitu dengan *web scraping*. *Web Scraping* ini memiliki beberapa langkah sebagai berikut:



Gambar 3. 3 Langkah-Langkah *Web Scraping*

HOME  
(HTTPS://LIBRARY.GUNADARMA.AC.ID/DEPC  
SYSTEM/)

PEMBAYARAN  
(HTTPS://LIBRARY.GUNADARMA.AC.ID/DEPC  
SYSTEM/PEMBAYARAN)

BIODATA DIRI  
(HTTPS://LIBRARY.GUNADARMA.AC.ID/DEPC  
SYSTEM/ANGGOTA)

PEMINJAMAN BUKU  
(HTTPS://LIBRARY.GUNADARMA.AC.ID/DEPC  
SYSTEM/ANGGOTA/PEMINJAMAN)

PENULISAN ILMIAH  
(HTTPS://LIBRARY.GUNADARMA.AC.ID/DEPC  
SYSTEM/ANGGOTA/PENULISAN)

FOTOCOPY  
(HTTPS://LIBRARY.GUNADARMA.AC.ID/DEPC  
SYSTEM/ANGGOTA/FOTOCOPY)

SUMBANG BUKU  
(HTTPS://LIBRARY.GUNADARMA.AC.ID/DEPC  
SYSTEM/ANGGOTA/SUMBANGBUKU)

KASUS & NOTIFIKASI  
(HTTPS://LIBRARY.GUNADARMA.AC.ID/DEPC  
SYSTEM/ANGGOTA/NOTIFIKASI)

BEBAS PERPUSTAKAAN  
(HTTPS://LIBRARY.GUNADARMA.AC.ID/DEPC  
SYSTEM/ANGGOTA/BEBASPERPUSTAKAAN)

KUNJUNGAN  
(HTTPS://LIBRARY.GUNADARMA.AC.ID/DEPC  
SYSTEM/KUNJUNGAN)

E-PAPER  
(HTTPS://LIBRARY.GUNADARMA.AC.ID/DEPC

Berdasarkan Gambar 3.3 di atas dapat dilihat bahwa langkah-langkah melakukan pengambilan data dengan *web scraping* sebagai berikut:

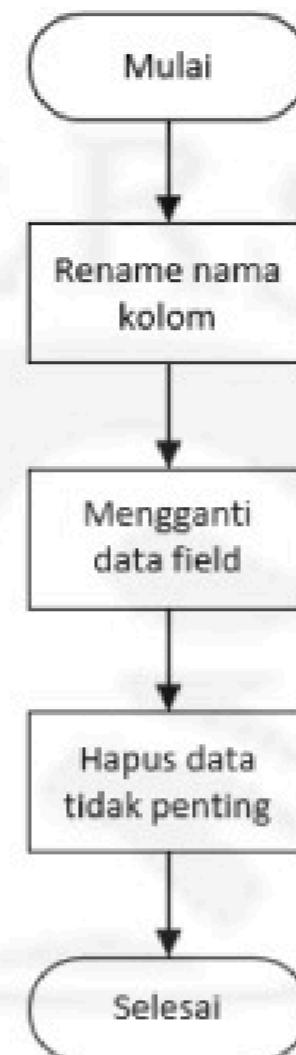
1. Teknik pengumpulan data diawali dengan menentukan URL dari *website* yang akan di *scrap*. Pada penelitian ini akan menggunakan adalah <https://www.rumah123.com/jual/jawa-barat/rumah/> sebagai *website* pengumpulan data pada provinsi Jawa Barat.
2. Setelah itu masukkan *user agent* yang ada pada bagian *network* saat *inspect website* tersebut.
3. Kemudian tentukan jumlah halaman yang akan di scrap pada *website* dengan sistem perulangan. Pengumpulan data yang dilakukan dari halaman 1 sampai 100 pada *website*.
4. Selanjutnya gunakan fungsi `requests.get` untuk mengakses URL tersebut, teks pada HTML akan disimpan dengan atribut `.text` dari *object* yang dikembalikan oleh `requests.get`.
5. Jika sudah lakukan parsing data HTML dengan *library* yang disediakan oleh *Python* yaitu *BeautifulSoup*.
6. Kemudian apabila sudah berada pada halaman akhir *scraping* tersebut maka akan disimpan ke dalam bentuk tabel dengan fungsi `csv.writer()`. Proses pengambilan data menggunakan *Web Scraping* sudah selesai.

*User agent* dapat dilihat pada bagian inspect lalu pada menu network. Jika respon yang diberikan *website* menghasilkan respon 200 maka artinya *website* tersebut sudah dapat di akses untuk mengambil data yang ada pada struktur HTML *website*. Perlu diperhatikan bahwa tidak semua *website* memberikan *user agent* jadi proses pemilihan *website* juga harus lebih teliti untuk melakukan *web scraping*. Diantara *website* yang memberikan pelayanan informasi penjualan rumah hanya pada rumah123.com yang menyediakan informasi lengkap mengenai kriteria yang akan digunakan pada penelitian ini dan respon 200 pada *user agent website*. Proses penyimpanan data ini menggunakan fungsi `csv.writer()` pada *Python* lalu menentukan nama kolom pada file tersebut.

- [HOME  
\(HTTPS://LIBRARY.GUNADARMA.AC.ID/DEPC SYSTEM/\)](#)
- [PEMBAYARAN  
\(HTTPS://LIBRARY.GUNADARMA.AC.ID/DEPC SYSTEM/PEMBAYARAN\)](#)
- [BIODATA DIRI  
\(HTTPS://LIBRARY.GUNADARMA.AC.ID/DEPC SYSTEM/ANGGOTA\)](#)
- [PEMINJAMAN BUKU  
\(HTTPS://LIBRARY.GUNADARMA.AC.ID/DEPC SYSTEM/ANGGOTA/PEMINJAMAN\)](#)
- [PENULISAN ILMIAH  
\(HTTPS://LIBRARY.GUNADARMA.AC.ID/DEPC SYSTEM/ANGGOTA/PENULISAN\)](#)
- [FOTOCOPY  
\(HTTPS://LIBRARY.GUNADARMA.AC.ID/DEPC SYSTEM/ANGGOTA/FOTOCOPY\)](#)
- [SUMBANG BUKU  
\(HTTPS://LIBRARY.GUNADARMA.AC.ID/DEPC SYSTEM/ANGGOTA/SUMBANGBUKU\)](#)
- [KASUS & NOTIFIKASI  
\(HTTPS://LIBRARY.GUNADARMA.AC.ID/DEPC SYSTEM/ANGGOTA/NOTIFIKASI\)](#)
- [BEBAS PERPUSTAKAAN  
\(HTTPS://LIBRARY.GUNADARMA.AC.ID/DEPC SYSTEM/ANGGOTA/BEBASPERPUSTAKAAN\)](#)
- [KUNJUNGAN  
\(HTTPS://LIBRARY.GUNADARMA.AC.ID/DEPC SYSTEM/KUNJUNGAN\)](#)
- [E-PAPER  
\(HTTPS://LIBRARY.GUNADARMA.AC.ID/DEPC SYSTEM/E-PAPER\)](#)

### 3.2.2. Data Validation

Pada proses ini akan menilai kesesuaian kualitas data dengan masalah yang akan dipecahkan.



Gambar 3. 4 Proses Data Validation

Berdasarkan Gambar 3.4 pada proses awal data validation adalah mengganti nama kolom pada data. Proses ini dapat dilakukan dengan manual yaitu mengganti nama kolom dengan Microsoft excel atau dengan *Python*.



Gambar 3. 5 Proses Perubahan Nama Kolom

Berdasarkan Gambar 3.5 di atas perubahan nama kolom ini dilakukan pada setiap kolom yang diperlukan saja. Perubahan nama kolom dengan menggunakan *Python* dapat mempercepat proses ini hanya dengan memanfaatkan fungsi `rename()` nama kolom yang sudah ditentukan akan berubah secara otomatis. Perubahan nama

HOME  
(HTTPS://LIBRARY.GUNADARMA.AC.ID/DEPC  
SYSTEM/)

PEMBAYARAN  
(HTTPS://LIBRARY.GUNADARMA.AC.ID/DEPC  
SYSTEM/PEMBAYARAN)

BIODATA DIRI  
(HTTPS://LIBRARY.GUNADARMA.AC.ID/DEPC  
SYSTEM/ANGGOTA)

PEMINJAMAN BUKU  
(HTTPS://LIBRARY.GUNADARMA.AC.ID/DEPC  
SYSTEM/ANGGOTA/PEMINJAMAN)

PENULISAN ILMIAH  
(HTTPS://LIBRARY.GUNADARMA.AC.ID/DEPC  
SYSTEM/ANGGOTA/PENULISAN)

FOTOCOPY  
(HTTPS://LIBRARY.GUNADARMA.AC.ID/DEPC  
SYSTEM/ANGGOTA/FOTOCOPY)

SUMBANG BUKU  
(HTTPS://LIBRARY.GUNADARMA.AC.ID/DEPC  
SYSTEM/ANGGOTA/SUMBANGBUKU)

KASUS & NOTIFIKASI  
(HTTPS://LIBRARY.GUNADARMA.AC.ID/DEPC  
SYSTEM/ANGGOTA/NOTIFIKASI)

BEBAS PERPUSTAKAAN  
(HTTPS://LIBRARY.GUNADARMA.AC.ID/DEPC  
SYSTEM/ANGGOTA/BEBASPERPUSTAKAAN)

KUNJUNGAN  
(HTTPS://LIBRARY.GUNADARMA.AC.ID/DEPC  
SYSTEM/KUNJUNGAN)

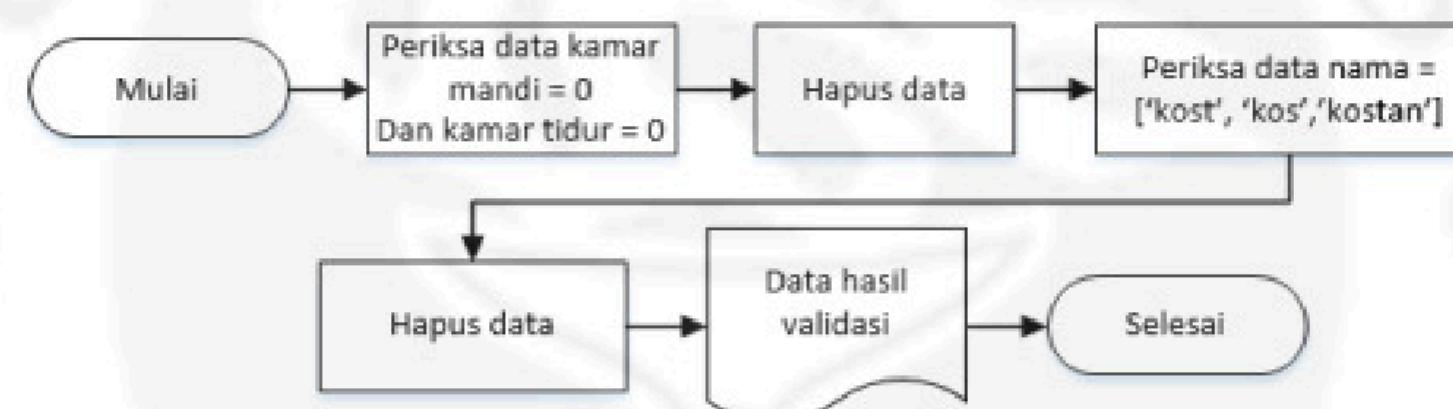
E-PAPER  
(HTTPS://LIBRARY.GUNADARMA.AC.ID/DEPC

kolom dapat memudahkan dalam proses pengolahan data untuk mengenal data yang digunakan.

Tabel 3. 4 Proses Perubahan Data Field

Data Asli	Proses	Setelah Diubah
500 Juta	replace ({' Juta' : '*1E6'}, regex =True)	500000000
1,9 Miliar	replace ({' Miliar' : '*1E9'}, regex =True)	19000000000

Berdasarkan Tabel 3.4 di atas adalah proses perubahan data *field* pada kolom harga yang akan mengubah tipe data *string* menjadi *integer*. Pada proses mengganti data *field* ini akan memudahkan proses pengolahan data dengan baik. Proses selanjutnya adalah pengecekan data yang tidak diperlukan.



Gambar 3. 6 Proses Pengecekan Data Tidak Diperlukan

Berdasarkan Gambar 3.6 di atas adalah proses pengecekan data yang tidak diperlukan pada kasus ini akan dicek data yang mempunyai jumlah tidak wajar. Pertama-tama dilakukan pemeriksaan terhadap data yang memiliki jumlah kamar mandi 0 dan kamar tidur 0 karena sangat tidak mungkin bahwa rumah tidak memiliki kamar mandi atau kamar tidur. Hapus data tersebut kemudian periksa data dengan nama rumah yang memiliki kata ‘kos’, ‘kostan’, ‘kost’ karena pada data yang akan diolah hanya berupa rumah sebagai tempat hidup standar sehingga rumah kost-kostan tidak diperlukan.

HOME  
(HTTPS://LIBRARY.GUNADARMA.AC.ID/DEPC  
SYSTEM/)

PEMBAYARAN  
(HTTPS://LIBRARY.GUNADARMA.AC.ID/DEPC  
SYSTEM/PEMBAYARAN)

BIODATA DIRI  
(HTTPS://LIBRARY.GUNADARMA.AC.ID/DEPC  
SYSTEM/ANGGOTA)

PEMINJAMAN BUKU  
(HTTPS://LIBRARY.GUNADARMA.AC.ID/DEPC  
SYSTEM/ANGGOTA/PEMINJAMAN)

PENULISAN ILMIAH  
(HTTPS://LIBRARY.GUNADARMA.AC.ID/DEPC  
SYSTEM/ANGGOTA/PENULISAN)

FOTOCOPY  
(HTTPS://LIBRARY.GUNADARMA.AC.ID/DEPC  
SYSTEM/ANGGOTA/FOTOCOPY)

SUMBANG BUKU  
(HTTPS://LIBRARY.GUNADARMA.AC.ID/DEPC  
SYSTEM/ANGGOTA/SUMBANGBUKU)

KASUS & NOTIFIKASI  
(HTTPS://LIBRARY.GUNADARMA.AC.ID/DEPC  
SYSTEM/ANGGOTA/NOTIFIKASI)

BEBAS PERPUSTAKAAN  
(HTTPS://LIBRARY.GUNADARMA.AC.ID/DEPC  
SYSTEM/ANGGOTA/BEBAСПERPUSTAKAAN)

KUNJUNGAN  
(HTTPS://LIBRARY.GUNADARMA.AC.ID/DEPC  
SYSTEM/KUNJUNGAN)

E-PAPER  
(HTTPS://LIBRARY.GUNADARMA.AC.ID/DEPC

### 3.2.3. Menelaah Data

Proses yang akan dilakukan pada tahap ini bertujuan untuk menganalisa data secara statistika deskriptif agar dapat melihat karakteristik dari setiap atribut.

Tabel 3. 5 Data Statistika Deskriptif

	Data statistika	Keterangan
Kolom numerik	<i>Count</i>	Jumlah dari semua data observasi
	<i>Sum</i>	Jumlah nilai dari semua data
	<i>Mean</i>	Nilai rata-rata pada data
	<i>Min</i>	Nilai minimum
	<i>Max</i>	Nilai maximum
	Kuartil 1	Nilai tengah antara nilai terkecil hingga data median
	Kuartil 2	Membagi data menjadi 2 (median)
	Kuartil 3	Nilai tengah antara median hingga nilai tertinggi
Kolom non-numerik	<i>Unique</i>	Banyaknya nilai unik dalam kolom
	<i>Top</i>	Nilai modus dari kolom
	<i>Freq</i>	Frekuensi dari nilai modus

Berdasarkan Tabel 3.5 di atas merupakan beberapa data statistik yang perlu di perhatikan saat mengecek statistika deskriptif pada data. Harus dilakukan pengecekan nilai korelasi antar kolom yang bertujuan untuk melihat keterhubungan kolom yang saling mempengaruhi.

Data yang akan diolah terdistribusi tidak normal dan menggunakan sumber data berbeda karena data rumah memiliki nilai kriteria yang berbeda, maka pada kasus ini menggunakan metode korelasi *Spearman* pada proses untuk melihat relasi keterhubungan antar kolom. Karena keterbatasan kemampuan yang dimiliki peneliti dalam mengolah data secara manual maka contoh perhitungan manual

HOME  
(HTTPS://LIBRARY.GUNADARMA.AC.ID/DEPC  
SYSTEM/)

PEMBAYARAN  
(HTTPS://LIBRARY.GUNADARMA.AC.ID/DEPC  
SYSTEM/PEMBAYARAN)

BIODATA DIRI  
(HTTPS://LIBRARY.GUNADARMA.AC.ID/DEPC  
SYSTEM/ANGGOTA)

PEMINJAMAN BUKU  
(HTTPS://LIBRARY.GUNADARMA.AC.ID/DEPC  
SYSTEM/ANGGOTA/PEMINJAMAN)

PENULISAN ILMIAH  
(HTTPS://LIBRARY.GUNADARMA.AC.ID/DEPC  
SYSTEM/ANGGOTA/PENULISAN)

FOTOCOPY  
(HTTPS://LIBRARY.GUNADARMA.AC.ID/DEPC  
SYSTEM/ANGGOTA/FOTOCOPY)

SUMBANG BUKU  
(HTTPS://LIBRARY.GUNADARMA.AC.ID/DEPC  
SYSTEM/ANGGOTA/SUMBANGBUKU)

KASUS & NOTIFIKASI  
(HTTPS://LIBRARY.GUNADARMA.AC.ID/DEPC  
SYSTEM/ANGGOTA/NOTIFIKASI)

BEBAS PERPUSTAKAAN  
(HTTPS://LIBRARY.GUNADARMA.AC.ID/DEPC  
SYSTEM/ANGGOTA/BEBASPERPUSTAKAAN)

KUNJUNGAN  
(HTTPS://LIBRARY.GUNADARMA.AC.ID/DEPC  
SYSTEM/KUNJUNGAN)

E-PAPER  
(HTTPS://LIBRARY.GUNADARMA.AC.ID/DEPC

korelasi *Spearman* berdasarkan Persamaan 2.5 akan menggunakan data sampel sebanyak 15 data.

Tabel 3. 6 Data Sampel Perhitungan Manual

Harga	Kamar tidur	Kamar mandi	Garasi	Luas tanah	Luas bangunan
3300000000	2	1	1	66	45
9000000000	3	1	1	75	108
12000000000	3	3	0	95	106
19500000000	3	3	2	111	200
14700000000	3	3	1	120	110
11000000000	3	2	1	93	104
2000000000	2	1	0	60	40
5850000000	2	1	2	75	45
268000000	2	2	1	60	30
1500000000	4	2	1	170	100
1010000000	3	2	0	89	103
8800000000	3	1	0	89	90
2300000000	1	2	2	63	61
1680000000	1	2	0	55	30
2100000000	4	3	0	190	180

Berdasarkan Tabel 3.6 di atas merupakan data sampel yang diambil dari 994 data hasil validasi kemudian akan digunakan untuk perhitungan manual. Proses perhitungan manual ini akan diberi contoh menghitung korelasi antara kolom luas bangunan dan luas tanah.

Tabel 3. 7 Perhitungan Korelasi *Spearman*

luas tanah (1)	luas bangunan (2)	rank 1	rank 2	d	d^2
66	45	11	11	0	0
75	108	9	4	5	25
95	106	5	5	0	0
111	200	4	1	3	9
120	110	3	3	0	0
93	104	6	6	0	0
60	40	13	13	0	0
75	45	9	11	2	4
60	30	13	14	1	1

HOME  
(HTTPS://LIBRARY.GUNADARMA.AC.ID/DEPC  
SYSTEM/)

PEMBAYARAN  
(HTTPS://LIBRARY.GUNADARMA.AC.ID/DEPC  
SYSTEM/PEMBAYARAN)

BIODATA DIRI  
(HTTPS://LIBRARY.GUNADARMA.AC.ID/DEPC  
SYSTEM/ANGGOTA)

PEMINJAMAN BUKU  
(HTTPS://LIBRARY.GUNADARMA.AC.ID/DEPC  
SYSTEM/ANGGOTA/PEMINJAMAN)

PENULISAN ILMIAH  
(HTTPS://LIBRARY.GUNADARMA.AC.ID/DEPC  
SYSTEM/ANGGOTA/PENULISAN)

FOTOCOPY  
(HTTPS://LIBRARY.GUNADARMA.AC.ID/DEPC  
SYSTEM/ANGGOTA/FOTOCOPY)

SUMBANG BUKU  
(HTTPS://LIBRARY.GUNADARMA.AC.ID/DEPC  
SYSTEM/ANGGOTA/SUMBANGBUKU)

KASUS & NOTIFIKASI  
(HTTPS://LIBRARY.GUNADARMA.AC.ID/DEPC  
SYSTEM/ANGGOTA/NOTIFIKASI)

BEBAS PERPUSTAKAAN  
(HTTPS://LIBRARY.GUNADARMA.AC.ID/DEPC  
SYSTEM/ANGGOTA/BEBAZPERPUSTAKAAN)

KUNJUNGAN  
(HTTPS://LIBRARY.GUNADARMA.AC.ID/DEPC  
SYSTEM/KUNJUNGAN)

E-PAPER  
(HTTPS://LIBRARY.GUNADARMA.AC.ID/DEPC

170	100	2	8	6	36
89	103	7	7	0	0
89	90	7	9	2	4
63	61	12	10	2	4
55	30	15	14	1	1
190	180	1	2	1	1

Berdasarkan Tabel 3.7 di atas merupakan hasil dari perhitungan  $d^2$ . Pada proses perhitungan ini pertama-tama lakukan pengurutan peringkat terhadap data korelasi pada contoh di atas menggunakan data luas tanah dan luas bangunan. Rank 1 akan menjelaskan peringkat dari data luas tanah sedangkan rank 2 akan menjelaskan peringkat pada data luas bangunan. Hitung nilai  $d$  yang merupakan hasil nilai *absolute* dari pengurangan nilai rank 1 dan rank 2. Kuadrat kan nilai tersebut untuk mendapat nilai akhir  $d^2$ .

Jika nilai  $d^2$  sudah didapatkan maka masukan hasil penjumlahan dari hasil  $d^2$  yaitu 85 ke dalam rumus korelasi.

$$\rho = 1 - \frac{6 \sum d_i^2}{n(n^2-1)}$$

Nilai  $n$  adalah jumlah data yang digunakan yaitu 15.

$$\rho = 1 - \frac{6 \times 85}{15(15^2-1)}$$

$$\rho = 1 - \frac{510}{3360}$$

$$\rho = 0.848214$$

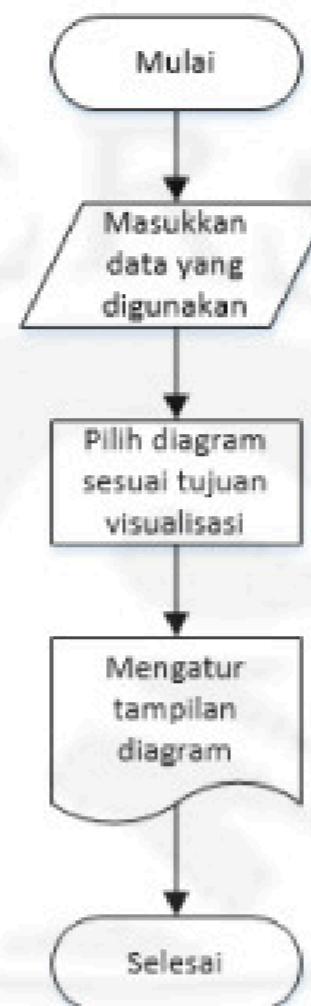
Hasil korelasi *Spearman* antara data luas tanah dan luas bangunan dengan 15 data sampel sebesar 0.84214. Lakukanlah hal tersebut pada semua pasangan dataset seperti luas bangunan dengan kamar tidur, luas bangunan dan kamar mandi, dan sebagainya sehingga mendapatkan hasil korelasi *Spearman* pada masing-masing kolom.

### 3.2.4. Eksplorasi Data dengan Visualisasi

Pada tahap ini akan dilakukan pemahaman data dengan metode visualisasi dalam bentuk diagram. Visualisasi memiliki peran yang cukup penting dalam implementasi *machine learning* dan *data science*. Tahap eksplorasi dengan

- [HOME  
\(HTTPS://LIBRARY.GUNADARMA.AC.ID/DEPC SYSTEM/\)](#)
- [PEMBAYARAN  
\(HTTPS://LIBRARY.GUNADARMA.AC.ID/DEPC SYSTEM/PEMBAYARAN\)](#)
- [BIODATA DIRI  
\(HTTPS://LIBRARY.GUNADARMA.AC.ID/DEPC SYSTEM/ANGGOTA\)](#)
- [PEMINJAMAN BUKU  
\(HTTPS://LIBRARY.GUNADARMA.AC.ID/DEPC SYSTEM/ANGGOTA/PEMINJAMAN\)](#)
- [PENULISAN ILMIAH  
\(HTTPS://LIBRARY.GUNADARMA.AC.ID/DEPC SYSTEM/ANGGOTA/PENULISAN\)](#)
- [FOTOCOPY  
\(HTTPS://LIBRARY.GUNADARMA.AC.ID/DEPC SYSTEM/ANGGOTA/FOTOCOPY\)](#)
- [SUMBANG BUKU  
\(HTTPS://LIBRARY.GUNADARMA.AC.ID/DEPC SYSTEM/ANGGOTA/SUMBANGBUKU\)](#)
- [KASUS & NOTIFIKASI  
\(HTTPS://LIBRARY.GUNADARMA.AC.ID/DEPC SYSTEM/ANGGOTA/NOTIFIKASI\)](#)
- [BEBAS PERPUSTAKAAN  
\(HTTPS://LIBRARY.GUNADARMA.AC.ID/DEPC SYSTEM/ANGGOTA/BEBASPERPUSTAKAAN\)](#)
- [KUNJUNGAN  
\(HTTPS://LIBRARY.GUNADARMA.AC.ID/DEPC SYSTEM/KUNJUNGAN\)](#)
- [E-PAPER  
\(HTTPS://LIBRARY.GUNADARMA.AC.ID/DEPC SYSTEM/E-PAPER\)](#)

visualisasi ini menceritakan tentang data yang akan di proses menggunakan tampilan yang mudah di mengerti oleh orang yang kurang berpengalaman dalam bidang data.



Gambar 3. 7 Proses Eksplorasi dengan Microsoft Power BI

Berdasarkan Gambar 3.7 di atas proses yang akan dilakukan pertama-tama adalah memasukkan data yang akan digunakan dan divisualisasi. Pilihlah diagram sesuai dengan tujuan dari visualisasi yang akan ditampilkan. Pada penelitian ini akan menggunakan diagram batang untuk memvisualisasikan data pada kolom kamar mandi dan kamar tidur. Tujuannya adalah untuk melihat frekuensi dari masing-masing jumlah kamar tidur dan kamat mandi yang dimiliki pada data rumah di provinsi Jawa Barat. Diagram lingkaran akan digunakan untuk melihat 5 provinsi terbanyak dengan rata-rata harga tertinggi. Tujuannya agar dapat memberikan informasi provinsi dengan nilai rata-rata harga rumah tertinggi dengan total rumah yang ada di provinsi tersebut. Diagram selanjutnya yang akan digunakan adalah penggabungan antara diagram garis dan diagram batang untuk visualisasi data garasi, luas tanah dan luas bangunan. Tujuannya agar dapat memperlihatkan

HOME  
(HTTPS://LIBRARY.GUNADARMA.AC.ID/DEPC  
SYSTEM/)

PEMBAYARAN  
(HTTPS://LIBRARY.GUNADARMA.AC.ID/DEPC  
SYSTEM/PEMBAYARAN)

BIODATA DIRI  
(HTTPS://LIBRARY.GUNADARMA.AC.ID/DEPC  
SYSTEM/ANGGOTA)

PEMINJAMAN BUKU  
(HTTPS://LIBRARY.GUNADARMA.AC.ID/DEPC  
SYSTEM/ANGGOTA/PEMINJAMAN)

PENULISAN ILMIAH  
(HTTPS://LIBRARY.GUNADARMA.AC.ID/DEPC  
SYSTEM/ANGGOTA/PENULISAN)

FOTOCOPY  
(HTTPS://LIBRARY.GUNADARMA.AC.ID/DEPC  
SYSTEM/ANGGOTA/FOTOCOPY)

SUMBANG BUKU  
(HTTPS://LIBRARY.GUNADARMA.AC.ID/DEPC  
SYSTEM/ANGGOTA/SUMBANGBUKU)

KASUS & NOTIFIKASI  
(HTTPS://LIBRARY.GUNADARMA.AC.ID/DEPC  
SYSTEM/ANGGOTA/NOTIFIKASI)

BEBAS PERPUSTAKAAN  
(HTTPS://LIBRARY.GUNADARMA.AC.ID/DEPC  
SYSTEM/ANGGOTA/BEBASPERPUSTAKAAN)

KUNJUNGAN  
(HTTPS://LIBRARY.GUNADARMA.AC.ID/DEPC  
SYSTEM/KUNJUNGAN)

E-PAPER  
(HTTPS://LIBRARY.GUNADARMA.AC.ID/DEPC

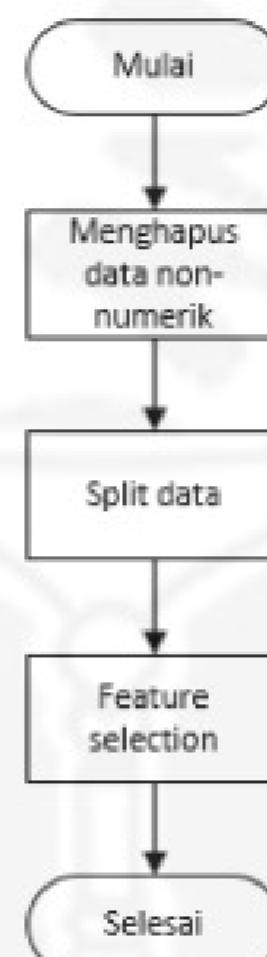
informasi seberapa besar pengaruhnya garasi pada rumah dengan luas tanah dan luas bangunan yang dimiliki.

### 3.3. Data Preparation

Pada tahap ini akan dilakukan perbaikan kualitas data yang sudah dilihat karakteristiknya sebelum masuk ke dalam tahap pemodelan.

#### 3.3.1. Seleksi data

Tahap seleksi data akan menentukan kolom yang memiliki pengaruh paling besar dengan target pada dataset.

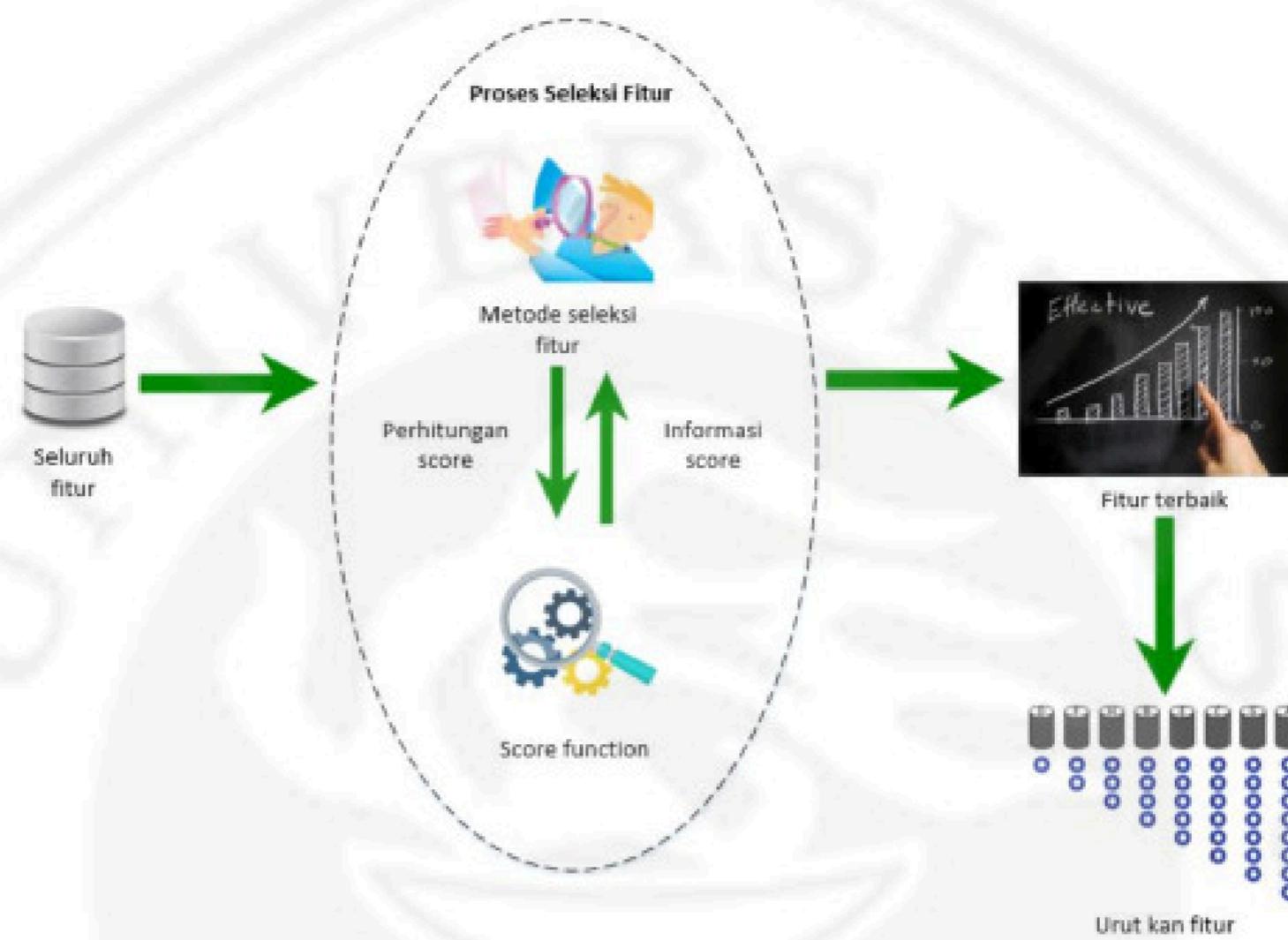


Gambar 3. 8 Proses Seleksi Data

Berdasarkan Gambar 3.8 di atas pertama-tama dalam proses seleksi data akan melakukan penghapusan data non-numerik karena proses pemilihan data akan dilakukan pada data numerik saja. Membagi data menjadi variabel *independent* (variabel yang mempengaruhi target) dan variabel target. Masukkan fungsi *feature selection* yang akan digunakan untuk mendapatkan fitur yang mempunyai pengaruh lebih banyak dengan target *output*. Pada *feature selection* dapat dipilih sesuai dengan masalah yang akan diselesaikan. Pemilihan metode *feature selection* akan

- [HOME  
\(HTTPS://LIBRARY.GUNADARMA.AC.ID/DEPC SYSTEM/\)](#)
- [PEMBAYARAN  
\(HTTPS://LIBRARY.GUNADARMA.AC.ID/DEPC SYSTEM/PEMBAYARAN\)](#)
- [BIODATA DIRI  
\(HTTPS://LIBRARY.GUNADARMA.AC.ID/DEPC SYSTEM/ANGGOTA\)](#)
- [PEMINJAMAN BUKU  
\(HTTPS://LIBRARY.GUNADARMA.AC.ID/DEPC SYSTEM/ANGGOTA/PEMINJAMAN\)](#)
- [PENULISAN ILMIAH  
\(HTTPS://LIBRARY.GUNADARMA.AC.ID/DEPC SYSTEM/ANGGOTA/PENULISAN\)](#)
- [FOTOCOPY  
\(HTTPS://LIBRARY.GUNADARMA.AC.ID/DEPC SYSTEM/ANGGOTA/FOTOCOPY\)](#)
- [SUMBANG BUKU  
\(HTTPS://LIBRARY.GUNADARMA.AC.ID/DEPC SYSTEM/ANGGOTA/SUMBANGBUKU\)](#)
- [KASUS & NOTIFIKASI  
\(HTTPS://LIBRARY.GUNADARMA.AC.ID/DEPC SYSTEM/ANGGOTA/NOTIFIKASI\)](#)
- [BEBAS PERPUSTAKAAN  
\(HTTPS://LIBRARY.GUNADARMA.AC.ID/DEPC SYSTEM/ANGGOTA/BEBASPERPUSTAKAAN\)](#)
- [KUNJUNGAN  
\(HTTPS://LIBRARY.GUNADARMA.AC.ID/DEPC SYSTEM/KUNJUNGAN\)](#)
- [E-PAPER  
\(HTTPS://LIBRARY.GUNADARMA.AC.ID/DEPC](#)

mempengaruhi nilai fitur yang di seleksi pada hasil akhirnya. *Feature selection* yang dapat digunakan pada masalah regresi adalah *selectbest*. Pada metode ini akan memfilter fitur sesuai dengan *score function* yang digunakan.



Gambar 3. 9 Proses Metode Selectbest

Berdasarkan Gambar 3.9 di atas *feature selection* dimulai dari seluruh fitur yang akan di seleksi. Proses selanjutnya adalah dari metode seleksi fitur yang digunakan yaitu *selectbest* kemudian akan menggunakan *score function* yaitu *f-regression* karena masalah yang akan diselesaikan berupa masalah regresi dalam memprediksi estimasi dari hasil data target. Metode *selectbest* akan menghitung *score* yang didapatkan pada masing-masing fitur menggunakan *f-regression*, lalu akan memberikan informasi *score* yang didapatkan. Fitur terbaik akan memiliki *score* dengan nilai paling tinggi kemudian fitur tersebut akan di urutkan sesuai dengan urutannya dimulai dari fitur paling tinggi ke paling rendah. Semakin tinggi nilai fitur maka semakin berpengaruh terhadap target yang dihasilkan.

HOME  
(HTTPS://LIBRARY.GUNADARMA.AC.ID/DEPC  
SYSTEM/)

PEMBAYARAN  
(HTTPS://LIBRARY.GUNADARMA.AC.ID/DEPC  
SYSTEM/PEMBAYARAN)

BIODATA DIRI  
(HTTPS://LIBRARY.GUNADARMA.AC.ID/DEPC  
SYSTEM/ANGGOTA)

PEMINJAMAN BUKU  
(HTTPS://LIBRARY.GUNADARMA.AC.ID/DEPC  
SYSTEM/ANGGOTA/PEMINJAMAN)

PENULISAN ILMIAH  
(HTTPS://LIBRARY.GUNADARMA.AC.ID/DEPC  
SYSTEM/ANGGOTA/PENULISAN)

FOTOCOPY  
(HTTPS://LIBRARY.GUNADARMA.AC.ID/DEPC  
SYSTEM/ANGGOTA/FOTOCOPY)

SUMBANG BUKU  
(HTTPS://LIBRARY.GUNADARMA.AC.ID/DEPC  
SYSTEM/ANGGOTA/SUMBANGBUKU)

KASUS & NOTIFIKASI  
(HTTPS://LIBRARY.GUNADARMA.AC.ID/DEPC  
SYSTEM/ANGGOTA/NOTIFIKASI)

BEBAS PERPUSTAKAAN  
(HTTPS://LIBRARY.GUNADARMA.AC.ID/DEPC  
SYSTEM/ANGGOTA/BEBASPERPUSTAKAAN)

KUNJUNGAN  
(HTTPS://LIBRARY.GUNADARMA.AC.ID/DEPC  
SYSTEM/KUNJUNGAN)

E-PAPER  
(HTTPS://LIBRARY.GUNADARMA.AC.ID/DEPC

### 3.3.2. Pembersihan Data

Pada tahap ini mempunyai tujuan untuk meminimalkan *noise*/data tidak lengkap/data salah sehingga data menjadi bersih dari bias/kesalahpahaman data.



Gambar 3. 10 Proses Pembersihan Data

Berdasarkan Gambar 3.10 di atas proses pembersihan data akan dilakukan dengan mengecek nilai kosong atau *missing value*. Beberapa metode yang dapat dilakukan untuk mengisi nilai kosong:

- Mengisi dengan nilai *mean*, jika *outlier* pada data berjumlah sedikit dan jumlah nilai kosong <5% dari jumlah data yang digunakan.
- Mengisi dengan nilai median, jika jumlah *outlier* pada data berjumlah banyak.
- Membiarkan saja, jika data dengan tipe *time series* sehingga sangat mempengaruhi hasil model.
- Menghapus 1 kolom, jika kolom tersebut memiliki nilai kosong > 50% dari jumlah data atau kolom tidak dibutuhkan.
- Mengisi dengan nilai suka-suka, jika probabilitas nilai kosong mungkin bergantung pada nilai yang diketahui.

HOME  
(HTTPS://LIBRARY.GUNADARMA.AC.ID/DEPC  
SYSTEM/)

PEMBAYARAN  
(HTTPS://LIBRARY.GUNADARMA.AC.ID/DEPC  
SYSTEM/PEMBAYARAN)

BIODATA DIRI  
(HTTPS://LIBRARY.GUNADARMA.AC.ID/DEPC  
SYSTEM/ANGGOTA)

PEMINJAMAN BUKU  
(HTTPS://LIBRARY.GUNADARMA.AC.ID/DEPC  
SYSTEM/ANGGOTA/PEMINJAMAN)

PENULISAN ILMIAH  
(HTTPS://LIBRARY.GUNADARMA.AC.ID/DEPC  
SYSTEM/ANGGOTA/PENULISAN)

FOTOCOPY  
(HTTPS://LIBRARY.GUNADARMA.AC.ID/DEPC  
SYSTEM/ANGGOTA/FOTOCOPY)

SUMBANG BUKU  
(HTTPS://LIBRARY.GUNADARMA.AC.ID/DEPC  
SYSTEM/ANGGOTA/SUMBANGBUKU)

KASUS & NOTIFIKASI  
(HTTPS://LIBRARY.GUNADARMA.AC.ID/DEPC  
SYSTEM/ANGGOTA/NOTIFIKASI)

BEBAS PERPUSTAKAAN  
(HTTPS://LIBRARY.GUNADARMA.AC.ID/DEPC  
SYSTEM/ANGGOTA/BEBASPERPUSTAKAAN)

KUNJUNGAN  
(HTTPS://LIBRARY.GUNADARMA.AC.ID/DEPC  
SYSTEM/KUNJUNGAN)

E-PAPER  
(HTTPS://LIBRARY.GUNADARMA.AC.ID/DEPC

- Mengisi dengan nilai modus, jika tipe data pada kolom nilai kosong adalah non-numerik dan mengalami *skew* pada data.
- Mengisi dengan teknik *Random Sample Imputer*, jika jenis sebaran nilai kosong adalah *Missing at Random* (MAR).

Harus di ketahui bahwa dari teknik yang dijelaskan tersebut tidak ada satu teknik yang sempurna dalam menangani nilai kosong pada data. Proses selanjutnya setelah mengetahui teknik untuk menangani nilai yang kosong maka lakukanlah pengecekan terhadap nilai duplikat. Satu-satunya jalan yang dapat dilakukan pada data yang terdeteksi duplikat adalah dengan menghapus data tersebut, lalu mengecek *outlier* pada data atau data pencilan.

Contoh perhitungan manual dalam mencari *outlier* menggunakan rumus IQR pada Persamaan 2.7 dengan 15 data sampel pada perhitungan sebelumnya.

Tabel 3. 8 Data Sampel Perhitungan *Outlier*

Luas tanah	luas tanah diurutkan
66	55
75	60
95	60
111	63
120	66
93	75
60	75
75	89
60	89
170	93
89	95
89	111
63	120
55	170
190	190

Berdasarkan Tabel 3.8 di atas merupakan 15 data sampel pada kolom luas tanah. Langkah pertama dalam mencari nilai kuartil adalah mengurutkan data dari yang terkecil hingga terbesar. Hitung nilai median (kuartil 2) dengan membagi data menjadi 2.

-  **HOME**  
[\(HTTPS://LIBRARY.GUNADARMA.AC.ID/DEPC SYSTEM/\)](https://library.gunadarma.ac.id/depc/system/)
  -  **PEMBAYARAN**  
[\(HTTPS://LIBRARY.GUNADARMA.AC.ID/DEPC SYSTEM/PEMBAYARAN\)](https://library.gunadarma.ac.id/depc/system/pembayaran)
  -  **BIODATA DIRI**  
[\(HTTPS://LIBRARY.GUNADARMA.AC.ID/DEPC SYSTEM/ANGGOTA\)](https://library.gunadarma.ac.id/depc/system/anggota)
  -  **PEMINJAMAN BUKU**  
[\(HTTPS://LIBRARY.GUNADARMA.AC.ID/DEPC SYSTEM/ANGGOTA/PEMINJAMAN\)](https://library.gunadarma.ac.id/depc/system/anggota/peminjaman)
  -  **PENULISAN ILMIAH**  
[\(HTTPS://LIBRARY.GUNADARMA.AC.ID/DEPC SYSTEM/ANGGOTA/PENULISAN\)](https://library.gunadarma.ac.id/depc/system/anggota/penulisan)
  -  **FOTOCOPY**  
[\(HTTPS://LIBRARY.GUNADARMA.AC.ID/DEPC SYSTEM/ANGGOTA/FOTOCOPY\)](https://library.gunadarma.ac.id/depc/system/anggota/fotocopy)
  -  **SUMBANG BUKU**  
[\(HTTPS://LIBRARY.GUNADARMA.AC.ID/DEPC SYSTEM/ANGGOTA/SUMBANGBUKU\)](https://library.gunadarma.ac.id/depc/system/anggota/sumbangbuku)
  -  **KASUS & NOTIFIKASI**  
[\(HTTPS://LIBRARY.GUNADARMA.AC.ID/DEPC SYSTEM/ANGGOTA/NOTIFIKASI\)](https://library.gunadarma.ac.id/depc/system/anggota/notifikasi)
  -  **BEBAS PERPUSTAKAAN**  
[\(HTTPS://LIBRARY.GUNADARMA.AC.ID/DEPC SYSTEM/ANGGOTA/BEBASPERPUSTAKAAN\)](https://library.gunadarma.ac.id/depc/system/anggota/bebasperpustakaan)
  -  **KUNJUNGAN**  
[\(HTTPS://LIBRARY.GUNADARMA.AC.ID/DEPC SYSTEM/KUNJUNGAN\)](https://library.gunadarma.ac.id/depc/system/kunjungan)
  -  **E-PAPER**  
[\(HTTPS://LIBRARY.GUNADARMA.AC.ID/DEPC E-PAPER\)](https://library.gunadarma.ac.id/depc/e-paper)

$$Q2 = 89$$

Dalam kasus ini nilai kuartil 2 didapatkan berdasarkan nilai pada urutan data ke-8 yaitu 75. Bagi data menjadi 2 bagian dengan arah ke atas dan ke bawah, data urutan 1-7 merupakan daerah data kuartil 1 dan data pada urutan 9-15 merupakan daerah data kuartil 3.

$$Q_1 = 63$$

$$Q3 = 111$$

$$IQR = Q3 - Q1 = 111 - 63 = 48$$

Syarat 1: Data  $< Q1 - 1.5 \times IQR$

Data < 63 – 1.5 × 48

Data < -9

Syarat 2: Data >  $Q3 + 1.5 \times IQR$

Data > 111 + 1.5 × 48

Data > 255

Berdasarkan contoh perhitungan IQR pada 15 data sampel di atas data sampel dikatakan *outlier* pada kolom luas tanah jika data  $< -9$  atau data  $> 255$ . Lakukanlah perhitungan tersebut kepada semua atribut seperti kolom harga, kamar tidur dan lainnya sehingga dapat mengecek data *outlier* pada 15 data sampel yang digunakan.

Tabel 3. 9 Hasil Perhitungan *Outlier*

HOME  
(HTTPS://LIBRARY.GUNADARMA.AC.ID/DEPC  
SYSTEM/)

PEMBAYARAN  
(HTTPS://LIBRARY.GUNADARMA.AC.ID/DEPC  
SYSTEM/PEMBAYARAN)

BIODATA DIRI  
(HTTPS://LIBRARY.GUNADARMA.AC.ID/DEPC  
SYSTEM/ANGGOTA)

PEMINJAMAN BUKU  
(HTTPS://LIBRARY.GUNADARMA.AC.ID/DEPC  
SYSTEM/ANGGOTA/PEMINJAMAN)

PENULISAN ILMIAH  
(HTTPS://LIBRARY.GUNADARMA.AC.ID/DEPC  
SYSTEM/ANGGOTA/PENULISAN)

FOTOCOPY  
(HTTPS://LIBRARY.GUNADARMA.AC.ID/DEPC  
SYSTEM/ANGGOTA/FOTOCOPY)

SUMBANG BUKU  
(HTTPS://LIBRARY.GUNADARMA.AC.ID/DEPC  
SYSTEM/ANGGOTA/SUMBANGBUKU)

KASUS & NOTIFIKASI  
(HTTPS://LIBRARY.GUNADARMA.AC.ID/DEPC  
SYSTEM/ANGGOTA/NOTIFIKASI)

BEBAS PERPUSTAKAAN  
(HTTPS://LIBRARY.GUNADARMA.AC.ID/DEPC  
SYSTEM/ANGGOTA/BEBASPERPUSTAKAAN)

KUNJUNGAN  
(HTTPS://LIBRARY.GUNADARMA.AC.ID/DEPC  
SYSTEM/KUNJUNGAN)

E-PAPER  
(HTTPS://LIBRARY.GUNADARMA.AC.ID/DEPC

FALSE	FALSE	FALSE	FALSE	FALSE	FALSE
FALSE	FALSE	FALSE	FALSE	FALSE	FALSE
FALSE	FALSE	FALSE	FALSE	FALSE	FALSE
FALSE	FALSE	FALSE	FALSE	FALSE	FALSE

Berdasarkan Tabel 3.9 di atas pada 15 data sampel yang digunakan untuk mengecek *outlier* menyatakan bahwa data sampel tidak memiliki *outlier* sesuai dengan hasil pada kolom *outlier* yaitu *false* yang berarti data yang digunakan tidak ada *outlier*. Proses dalam pengecekan *outlier* selain dengan perhitungan IQR dapat menggunakan dengan menghitung nilai *Z-Score* pada data dan menggunakan visualisasi *boxplot*.

### 3.3.3. Konstruksi Data

Proses selanjutnya setelah melakukan pembersihan data maka adalah konstruksi data untuk menambahkan fitur, pelabelan data dan melakukan transformasi data non-numerik menjadi numerik. Kuantitas dan kualitas suatu data secara langsung dapat menentukan keberhasilan suatu algoritma *machine learning*. Hal ini yang menjadi sebuah alasan mengapa proses pelabelan pada data merupakan bagian yang paling penting dari alur kerja persiapan data dalam membangun model *machine learning* yang baik. Proses transformasi data ini dengan menggunakan labeling pada data unik seperti nilai unik untuk data kota bandung akan diberi label 1 untuk hasil pelabelan data.

- [HOME  
\(HTTPS://LIBRARY.GUNADARMA.AC.ID/DEPC SYSTEM/\)](#)
- [PEMBAYARAN  
\(HTTPS://LIBRARY.GUNADARMA.AC.ID/DEPC SYSTEM/PEMBAYARAN\)](#)
- [BIODATA DIRI  
\(HTTPS://LIBRARY.GUNADARMA.AC.ID/DEPC SYSTEM/ANGGOTA\)](#)
- [PEMINJAMAN BUKU  
\(HTTPS://LIBRARY.GUNADARMA.AC.ID/DEPC SYSTEM/ANGGOTA/PEMINJAMAN\)](#)
- [PENULISAN ILMIAH  
\(HTTPS://LIBRARY.GUNADARMA.AC.ID/DEPC SYSTEM/ANGGOTA/PENULISAN\)](#)
- [FOTOCOPY  
\(HTTPS://LIBRARY.GUNADARMA.AC.ID/DEPC SYSTEM/ANGGOTA/FOTOCOPY\)](#)
- [SUMBANG BUKU  
\(HTTPS://LIBRARY.GUNADARMA.AC.ID/DEPC SYSTEM/ANGGOTA/SUMBANGBUKU\)](#)
- [KASUS & NOTIFIKASI  
\(HTTPS://LIBRARY.GUNADARMA.AC.ID/DEPC SYSTEM/ANGGOTA/NOTIFIKASI\)](#)
- [BEBAS PERPUSTAKAAN  
\(HTTPS://LIBRARY.GUNADARMA.AC.ID/DEPC SYSTEM/ANGGOTA/BEBAisperpustakaan\)](#)
- [KUNJUNGAN  
\(HTTPS://LIBRARY.GUNADARMA.AC.ID/DEPC SYSTEM/KUNJUNGAN\)](#)
- [E-PAPER  
\(HTTPS://LIBRARY.GUNADARMA.AC.ID/DEPC](#)



Gambar 3. 11 Pelabelan Data *Label encoding*

Berdasarkan Gambar 3.11 di atas proses pelabelan data yang akan digunakan pada penelitian ini adalah menggunakan teknik *label encoding* karena pada teknik ini akan mengutamakan urutan dari data sehingga label yang akan dihasilkan berupa nilai bilangan bulan yang berurutan. Pertama-tama data non-numerik akan diurutkan berdasarkan abjad, kemudian proses pelabelan dengan teknik *label encoding* akan mengkonversi nilai unik pada data non-numerik menjadi *integer* sesuai dengan urutannya. Proses konversi ini tidak membuat kolom baru sehingga kolom pada data non-numerik akan berubah secara otomatis menjadi tipe data *integer*. Pelabelan data akan di awali dengan nilai angka bulat 0 dan diikuti penambahan nilai sesuai dengan jumlah nilai unik.

Pada tahap konstruksi juga dilakukan normalisasi data atau yang biasa disebut sebagai penyamarataan rentang pada data. Dari data penelitian ini dapat dilihat bahwa rentang data harga dengan data lainnya memiliki rentang yang cukup jauh, maka dari itu perlu dilakukan normalisasi data agar rentang semua data latih sebelum di masukan ke dalam model memiliki rentang yang sama.

HOME  
(HTTPS://LIBRARY.GUNADARMA.AC.ID/DEPC  
SYSTEM/)

PEMBAYARAN  
(HTTPS://LIBRARY.GUNADARMA.AC.ID/DEPC  
SYSTEM/PEMBAYARAN)

BIODATA DIRI  
(HTTPS://LIBRARY.GUNADARMA.AC.ID/DEPC  
SYSTEM/ANGGOTA)

PEMINJAMAN BUKU  
(HTTPS://LIBRARY.GUNADARMA.AC.ID/DEPC  
SYSTEM/ANGGOTA/PEMINJAMAN)

PENULISAN ILMIAH  
(HTTPS://LIBRARY.GUNADARMA.AC.ID/DEPC  
SYSTEM/ANGGOTA/PENULISAN)

FOTOCOPY  
(HTTPS://LIBRARY.GUNADARMA.AC.ID/DEPC  
SYSTEM/ANGGOTA/FOTOCOPY)

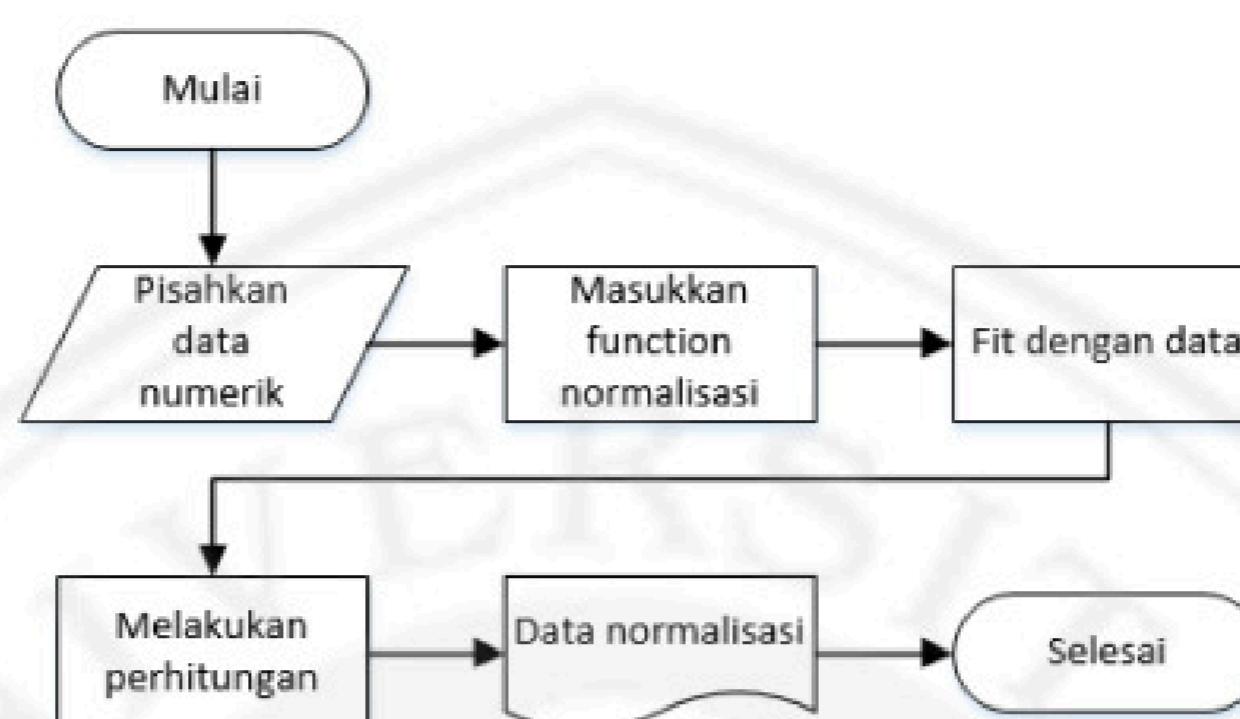
SUMBANG BUKU  
(HTTPS://LIBRARY.GUNADARMA.AC.ID/DEPC  
SYSTEM/ANGGOTA/SUMBANGBUKU)

KASUS & NOTIFIKASI  
(HTTPS://LIBRARY.GUNADARMA.AC.ID/DEPC  
SYSTEM/ANGGOTA/NOTIFIKASI)

BEBAS PERPUSTAKAAN  
(HTTPS://LIBRARY.GUNADARMA.AC.ID/DEPC  
SYSTEM/ANGGOTA/BEBASPERPUSTAKAAN)

KUNJUNGAN  
(HTTPS://LIBRARY.GUNADARMA.AC.ID/DEPC  
SYSTEM/KUNJUNGAN)

E-PAPER  
(HTTPS://LIBRARY.GUNADARMA.AC.ID/DEPC



Gambar 3. 12 Normalisasi data

Berdasarkan Gambar 3.12 di atas normalisasi data digunakan hanya untuk data numerik maka dari itu dilakukan pemisahan terhadap data non-numerik. Dikarenakan pada data penelitian tidak terdistribusi normal dan proses pembersihan data sudah dilakukan penghapusan data *outlier* maka *function* normalisasi yang digunakan pada penelitian ini akan menggunakan metode *MinMaxScaler* dikarenakan metode ini tidak menyebabkan distorsi pada data atau penyimpangan data, jadi hasil yang diberikan akan mempertahankan bentuk data. Data hasil perhitungan normalisasi akan membentuk suatu data dengan rentang 0-1 sebagai data yang sudah di normalisasi.

Pada proses perhitungan akan menggunakan Persamaan 2.10 dengan 15 data sampel yang telah digunakan pada perhitungan sebelumnya.

Tabel 3. 10 Data Sebelum Normalisasi

	daerah	kamar tidur	kamar mandi	garasi	luas tanah	luas bangunan	harga
	8	2	1	1	66	45	268000000
	2	3	1	1	75	108	900000000
	0	3	3	0	95	106	1200000000
	0	3	3	2	111	200	1950000000
	8	3	3	1	120	110	1470000000
	0	3	2	1	93	104	1100000000
	0	2	1	0	60	40	2000000000

HOME  
(HTTPS://LIBRARY.GUNADARMA.AC.ID/DEPC  
SYSTEM/)

PEMBAYARAN  
(HTTPS://LIBRARY.GUNADARMA.AC.ID/DEPC  
SYSTEM/PEMBAYARAN)

BIODATA DIRI  
(HTTPS://LIBRARY.GUNADARMA.AC.ID/DEPC  
SYSTEM/ANGGOTA)

PEMINJAMAN BUKU  
(HTTPS://LIBRARY.GUNADARMA.AC.ID/DEPC  
SYSTEM/ANGGOTA/PEMINJAMAN)

PENULISAN ILMIAH  
(HTTPS://LIBRARY.GUNADARMA.AC.ID/DEPC  
SYSTEM/ANGGOTA/PENULISAN)

FOTOCOPY  
(HTTPS://LIBRARY.GUNADARMA.AC.ID/DEPC  
SYSTEM/ANGGOTA/FOTOCOPY)

SUMBANG BUKU  
(HTTPS://LIBRARY.GUNADARMA.AC.ID/DEPC  
SYSTEM/ANGGOTA/SUMBANGBUKU)

KASUS & NOTIFIKASI  
(HTTPS://LIBRARY.GUNADARMA.AC.ID/DEPC  
SYSTEM/ANGGOTA/NOTIFIKASI)

BEBAS PERPUSTAKAAN  
(HTTPS://LIBRARY.GUNADARMA.AC.ID/DEPC  
SYSTEM/ANGGOTA/BEBASPERPUSTAKAAN)

KUNJUNGAN  
(HTTPS://LIBRARY.GUNADARMA.AC.ID/DEPC  
SYSTEM/KUNJUNGAN)

E-PAPER  
(HTTPS://LIBRARY.GUNADARMA.AC.ID/DEPC

	8	2	1	2	75	45	585000000
	3	2	2	1	60	30	330000000
	2	4	2	1	170	100	1500000000
	8	3	2	0	89	103	1010000000
	8	3	1	0	89	90	880000000
	8	1	2	2	63	61	230000000
	2	1	2	0	55	30	168000000
	0	4	3	0	190	180	2100000000
min	0	1	1	0	55	30	168000000
max	8	4	3	2	190	200	2100000000

Berdasarkan Tabel 3.10 di atas adalah data sampel sebelum dilakukan normalisasi. Pada tabel tersebut sudah ditentukan nilai minimal dan maksimal dari setiap kolom, langkah selanjutnya adalah memasukkan angka tersebut ke dalam Persamaan 2.10 dengan nilai newmin = 0 dan newmax = 1.

$$\text{newdata} = \frac{(data - \text{min}) \times (\text{newmax} - \text{newmin})}{(\text{max} - \text{min})} + \text{newmin}$$

$$\text{newdatadaerah}(1) = \frac{(8-0) \times (1-0)}{(8-0)} + 0$$

$$\text{newdatadaerah}(1) = \frac{8}{8} + 0 = 1$$

Perhitungan ini dilakukan dengan semua data yang ada pada kolom data sampel. Hasil yang akan diberikan berupa data normalisasi dengan rentang nilai 0 sampai 1.

Tabel 3. 11 Data Sesudah Normalisasi

daerah	kamar tidur	kamar mandi	garasi	luas tanah	luas bangunan	harga
1	0.333333	0	0.5	0.081481	0.088235	0.05176
0.25	0.666667	0	0.5	0.148148	0.458824	0.378882
0	0.666667	1	0	0.296296	0.447059	0.534161
0	0.666667	1	1	0.414815	1	0.92236
1	0.666667	1	0.5	0.481481	0.470588	0.673913
0	0.666667	0.5	0.5	0.281481	0.435294	0.482402
0	0.333333	0	0	0.037037	0.058824	0.016563
1	0.333333	0	1	0.148148	0.088235	0.215839
0.375	0.333333	0.5	0.5	0.037037	0	0.083851

HOME  
(HTTPS://LIBRARY.GUNADARMA.AC.ID/DEPC  
SYSTEM/)

PEMBAYARAN  
(HTTPS://LIBRARY.GUNADARMA.AC.ID/DEPC  
SYSTEM/PEMBAYARAN)

BIODATA DIRI  
(HTTPS://LIBRARY.GUNADARMA.AC.ID/DEPC  
SYSTEM/ANGGOTA)

PEMINJAMAN BUKU  
(HTTPS://LIBRARY.GUNADARMA.AC.ID/DEPC  
SYSTEM/ANGGOTA/PEMINJAMAN)

PENULISAN ILMIAH  
(HTTPS://LIBRARY.GUNADARMA.AC.ID/DEPC  
SYSTEM/ANGGOTA/PENULISAN)

FOTOCOPY  
(HTTPS://LIBRARY.GUNADARMA.AC.ID/DEPC  
SYSTEM/ANGGOTA/FOTOCOPY)

SUMBANG BUKU  
(HTTPS://LIBRARY.GUNADARMA.AC.ID/DEPC  
SYSTEM/ANGGOTA/SUMBANGBUKU)

KASUS & NOTIFIKASI  
(HTTPS://LIBRARY.GUNADARMA.AC.ID/DEPC  
SYSTEM/ANGGOTA/NOTIFIKASI)

BEBAS PERPUSTAKAAN  
(HTTPS://LIBRARY.GUNADARMA.AC.ID/DEPC  
SYSTEM/ANGGOTA/BEBASTERPUSTAKAAN)

KUNJUNGAN  
(HTTPS://LIBRARY.GUNADARMA.AC.ID/DEPC  
SYSTEM/KUNJUNGAN)

E-PAPER  
(HTTPS://LIBRARY.GUNADARMA.AC.ID/DEPC

0.25	1	0.5	0.5	0.851852	0.411765	0.689441
1	0.666667	0.5	0	0.251852	0.429412	0.435818
1	0.666667	0	0	0.251852	0.352941	0.36853
1	0	0.5	1	0.059259	0.182353	0.032091
0.25	0	0.5	0	0	0	0
0	1	1	0	1	0.882353	1

Berdasarkan Tabel 3.11 di atas dapat dilihat bahwa hasil data sesudah normalisasi memiliki rentang yang sama pada setiap kolom yaitu 0-1. Data hasil normalisasi ini merupakan data yang sudah siap untuk dilakukan pemodelan dengan algoritma *machine learning* pada tahap selanjutnya.

### 3.4. Modeling

Pada proses ini akan dilakukan pemodelan data latih dengan algoritma *Random Forest Regression* untuk memprediksi harga rumah dan *Cosine Similarity* untuk memberikan rekomendasi rumah berdasarkan hasil prediksi.

#### 3.4.1. Pemisahan Data

Pada tahap pemisahan data ini akan dilakukan menjadi 3 kriteria yaitu sebagai berikut:

Tabel 3. 12 Kriteria Pemisahan Data

Rasio	Data	Persentase
80:20	Latih	80%
	Uji	20%
75:25	Latih	75%
	Uji	25%
70:30	Latih	70%
	Uji	30%

Berdasarkan Tabel 3.12 di atas berisi 3 kriteria pada masing-masing kriteria akan diambil data secara acak untuk data latih dan data uji. Data latih akan berfungsi untuk melakukan pelatihan data terhadap model yang akan digunakan, sedangkan

HOME  
(HTTPS://LIBRARY.GUNADARMA.AC.ID/DEPC  
SYSTEM/)

PEMBAYARAN  
(HTTPS://LIBRARY.GUNADARMA.AC.ID/DEPC  
SYSTEM/PEMBAYARAN)

BIODATA DIRI  
(HTTPS://LIBRARY.GUNADARMA.AC.ID/DEPC  
SYSTEM/ANGGOTA)

PEMINJAMAN BUKU  
(HTTPS://LIBRARY.GUNADARMA.AC.ID/DEPC  
SYSTEM/ANGGOTA/PEMINJAMAN)

PENULISAN ILMIAH  
(HTTPS://LIBRARY.GUNADARMA.AC.ID/DEPC  
SYSTEM/ANGGOTA/PENULISAN)

FOTOCOPY  
(HTTPS://LIBRARY.GUNADARMA.AC.ID/DEPC  
SYSTEM/ANGGOTA/FOTOCOPY)

SUMBANG BUKU  
(HTTPS://LIBRARY.GUNADARMA.AC.ID/DEPC  
SYSTEM/ANGGOTA/SUMBANGBUKU)

KASUS & NOTIFIKASI  
(HTTPS://LIBRARY.GUNADARMA.AC.ID/DEPC  
SYSTEM/ANGGOTA/NOTIFIKASI)

BEBAS PERPUSTAKAAN  
(HTTPS://LIBRARY.GUNADARMA.AC.ID/DEPC  
SYSTEM/ANGGOTA/BEBASTERPUSTAKAAN)

KUNJUNGAN  
(HTTPS://LIBRARY.GUNADARMA.AC.ID/DEPC  
SYSTEM/KUNJUNGAN)

E-PAPER  
(HTTPS://LIBRARY.GUNADARMA.AC.ID/DEPC

data uji akan berfungsi untuk melakukan evaluasi terhadap model dengan data yang belum dikenali oleh model. Jadi data uji akan dianggap sebagai data baru yang akan mempelajari model sehingga mendapatkan nilai akurasi dari proses pengujian tersebut. Dari 3 kriteria tersebut akan dipilih kriteria pemisahan data dengan pelatihan model yang mendapatkan hasil evaluasi prediksi paling tinggi akan dijadikan sebagai model untuk proses *deployment*.

### 3.4.2. Modeling Algoritma Random Forest Regression

Proses *modeling* ini akan dilakukan dengan 2 cara yaitu menggunakan bahasa pemrograman *Python* dan menggunakan perhitungan manual sebagai pembuktian dari sistem. Pada proses *modeling* karena adanya keterbatasan kemampuan yang dimiliki penulis maka pada proses perhitungan manual akan menggunakan 15 data sampel *bootstrap* yang dipilih secara acak pada dataset.

Tabel 3. 13 Data Sampel *Bootstrap* Perhitungan Random Forest Regression

Predictors						Target
daerah	kamar tidur	kamar mandi	garasi	luas tanah	luas bangunan	harga
8	2	1	1	66	45	2680000000
2	3	1	1	75	108	9000000000
0	3	3	0	95	106	1200000000
0	3	3	2	111	200	1950000000
8	3	3	1	120	110	1470000000
0	3	2	1	93	104	1100000000
0	2	1	0	60	40	2000000000
8	2	1	2	75	45	585000000
3	2	2	1	60	30	330000000
2	4	2	1	170	100	1500000000
8	3	2	0	89	103	1010000000
8	3	1	0	89	90	880000000
8	1	2	2	63	61	230000000
2	1	2	0	55	30	168000000
0	4	3	0	190	180	2100000000

HOME  
(HTTPS://LIBRARY.GUNADARMA.AC.ID/DEPC  
SYSTEM/)

PEMBAYARAN  
(HTTPS://LIBRARY.GUNADARMA.AC.ID/DEPC  
SYSTEM/PEMBAYARAN)

BIODATA DIRI  
(HTTPS://LIBRARY.GUNADARMA.AC.ID/DEPC  
SYSTEM/ANGGOTA)

PEMINJAMAN BUKU  
(HTTPS://LIBRARY.GUNADARMA.AC.ID/DEPC  
SYSTEM/ANGGOTA/PEMINJAMAN)

PENULISAN ILMIAH  
(HTTPS://LIBRARY.GUNADARMA.AC.ID/DEPC  
SYSTEM/ANGGOTA/PENULISAN)

FOTOCOPY  
(HTTPS://LIBRARY.GUNADARMA.AC.ID/DEPC  
SYSTEM/ANGGOTA/FOTOCOPY)

SUMBANG BUKU  
(HTTPS://LIBRARY.GUNADARMA.AC.ID/DEPC  
SYSTEM/ANGGOTA/SUMBANGBUKU)

KASUS & NOTIFIKASI  
(HTTPS://LIBRARY.GUNADARMA.AC.ID/DEPC  
SYSTEM/ANGGOTA/NOTIFIKASI)

BEBAS PERPUSTAKAAN  
(HTTPS://LIBRARY.GUNADARMA.AC.ID/DEPC  
SYSTEM/ANGGOTA/BEBASPERPUSTAKAAN)

KUNJUNGAN  
(HTTPS://LIBRARY.GUNADARMA.AC.ID/DEPC  
SYSTEM/KUNJUNGAN)

E-PAPER  
(HTTPS://LIBRARY.GUNADARMA.AC.ID/DEPC

Berdasarkan Tabel 3.13 di atas merupakan data sampel *bootstrap* yang akan digunakan untuk proses contoh perhitungan manual dengan *Random Forest Regression*.

Tabel 3. 14 Data Seleksi Variabel Acak

kamar tidur	kamar mandi	garasi	harga
2	1	1	2680000000
3	1	1	9000000000
3	3	0	12000000000
3	3	2	19500000000
3	3	1	14700000000
3	2	1	11000000000
2	1	0	2000000000
2	1	2	5850000000
2	2	1	3300000000
4	2	1	15000000000
3	2	0	10100000000
3	1	0	8800000000
1	2	2	2300000000
1	2	0	1680000000
4	3	0	2100000000

Berdasarkan Tabel 3.14 di atas merupakan tabel hasil pemilihan variabel *predictors* secara acak dengan target variabel harga rumah. Data ini akan digunakan untuk pembentukan *Random Forest* dengan kumpulan pohon *Decision Tree*.

Pertama hitung nilai *weight average MSE* pada kolom kamar tidur. Pada perhitungan ini akan dimulai dengan menghitung probabilitas dari masing-masing jumlah kamar kemudian menghitung rata-rata harga dari masing-masing jumlah kamar tidur.

$$\text{PROP KAMAR TIDUR 1: } \frac{2}{15} = 0.13333$$

$$\text{PROP KAMAR TIDUR 2: } \frac{4}{15} = 0.26667$$

$$\text{PROP KAMAR TIDUR 3: } \frac{7}{15} = 0.46667$$

$$\text{PROP KAMAR TIDUR 4: } \frac{2}{15} = 0.13333$$

$$\text{AVG KAMAR TIDUR 1: } \frac{168000000 + 230000000}{2} = 199000000$$

HOME  
(HTTPS://LIBRARY.GUNADARMA.AC.ID/DEPC  
SYSTEM/)

PEMBAYARAN  
(HTTPS://LIBRARY.GUNADARMA.AC.ID/DEPC  
SYSTEM/PEMBAYARAN)

BIODATA DIRI  
(HTTPS://LIBRARY.GUNADARMA.AC.ID/DEPC  
SYSTEM/ANGGOTA)

PEMINJAMAN BUKU  
(HTTPS://LIBRARY.GUNADARMA.AC.ID/DEPC  
SYSTEM/ANGGOTA/PEMINJAMAN)

PENULISAN ILMIAH  
(HTTPS://LIBRARY.GUNADARMA.AC.ID/DEPC  
SYSTEM/ANGGOTA/PENULISAN)

FOTOCOPY  
(HTTPS://LIBRARY.GUNADARMA.AC.ID/DEPC  
SYSTEM/ANGGOTA/FOTOCOPY)

SUMBANG BUKU  
(HTTPS://LIBRARY.GUNADARMA.AC.ID/DEPC  
SYSTEM/ANGGOTA/SUMBANGBUKU)

KASUS & NOTIFIKASI  
(HTTPS://LIBRARY.GUNADARMA.AC.ID/DEPC  
SYSTEM/ANGGOTA/NOTIFIKASI)

BEBAS PERPUSTAKAAN  
(HTTPS://LIBRARY.GUNADARMA.AC.ID/DEPC  
SYSTEM/ANGGOTA/BEBAСПERPUSTAKAAN)

KUNJUNGAN  
(HTTPS://LIBRARY.GUNADARMA.AC.ID/DEPC  
SYSTEM/KUNJUNGAN)

E-PAPER  
(HTTPS://LIBRARY.GUNADARMA.AC.ID/DEPC

$$\text{AVG KAMAR TIDUR 2: } \frac{3300000000 + 200000000 + 585000000 + 268000000}{4} =$$

$$345750000$$

AVG KAMAR TIDUR 3:

$$\frac{900000000 + 1200000000 + 1950000000 + 1470000000 + 1100000000 + 1010000000 + 880000000}{7} =$$

$$1345555556$$

$$\text{AVG KAMAR TIDUR 4: } \frac{1500000000 + 2100000000}{2} = 1800000000$$

Jika sudah maka hitung masing-masing nilai MSE pada setiap jumlah kamar tidur berdasarkan persamaan 2.1.

$$\begin{aligned} \text{MSE KAMAR TIDUR 1: } & \frac{(168000000 - 199000000)^2}{2} + \frac{(230000000 - 199000000)^2}{2} \\ & = (480.500.000.000.000) + (480.500.000.000.000) \\ & = 9.61e + 14 \end{aligned}$$

Hal ini dilakukan pada semua jumlah kamar tidur sehingga mendapatkan hasil MSE masing-masing sebagai berikut:

$$\text{MSE KAMAR TIDUR 2: } 2.11942E+16$$

$$\text{MSE KAMAR TIDUR 3: } 1.41455E+17$$

$$\text{MSE KAMAR TIDUR 4: } 9E+16$$

Hitung nilai *weight average* MSE dengan hasil penjumlahan dari perkalian masing-masing MSE dengan nilai probabilitas.

$$\begin{aligned} \text{Weight Average MSE KAMAR TIDUR: } & ((9.61e + 14) \times 0.13333) + \\ & ((2.11942E + 16) \times 0.26667) + ((1.41455E + 17) \times 0.46667) + ((9e + 16) \times 0.13333) = 8.37921e + 16 \end{aligned}$$

Semua langkah di atas diulangi untuk mendapatkan nilai *weight average* MSE pada kolom kamar mandi dan garasi. Sehingga akan mendapatkan variabel untuk melakukan *splitting* pada pohon yang akan dibentuk.

Tabel 3. 15 Hasil Perhitungan MSE Terkecil dari 3 Variabel

KAMAR TIDUR	MSE (HARGA)	COUNT	PROP	AVG
1	9.61E+14	2	0.133333333	199000000
2	2.11942E+16	4	0.266666667	345750000
3	1.41455E+17	7	0.466666667	1345555556

[HOME  
\(HTTPS://LIBRARY.GUNADARMA.AC.ID/DEPC  
SYSTEM/\)](https://library.gunadarma.ac.id/DEPC/system/)

[\*\*PEMBAYARAN\*\*  
\(HTTPS://LIBRARY.GUNADARMA.AC.ID/DEPC  
SYSTEM/PEMBAYARAN\)](https://library.gunadarma.ac.id/DEPC/system/pembayaran)

[\*\*BIODATA DIRI\*\*  
\(HTTPS://LIBRARY.GUNADARMA.AC.ID/DEPC  
SYSTEM/ANGGOTA\)](https://library.gunadarma.ac.id/DEPC/system/anggota/biodata)

[\*\*PEMINJAMAN BUKU\*\*  
\(HTTPS://LIBRARY.GUNADARMA.AC.ID/DEPC  
SYSTEM/ANGGOTA/PEMINJAMAN\)](https://library.gunadarma.ac.id/DEPC/system/anggota/peminjaman)

[\*\*PENULISAN ILMIAH\*\*  
\(HTTPS://LIBRARY.GUNADARMA.AC.ID/DEPC  
SYSTEM/ANGGOTA/PENULISAN\)](https://library.gunadarma.ac.id/DEPC/system/anggota/penulisan)

[\*\*FOTOCOPY\*\*  
\(HTTPS://LIBRARY.GUNADARMA.AC.ID/DEPC  
SYSTEM/ANGGOTA/FOTOCOPY\)](https://library.gunadarma.ac.id/DEPC/system/anggota/fotocopy)

[\*\*SUMBANG BUKU\*\*  
\(HTTPS://LIBRARY.GUNADARMA.AC.ID/DEPC  
SYSTEM/ANGGOTA/SUMBANGBUKU\)](https://library.gunadarma.ac.id/DEPC/system/anggota/sumbang)

[\*\*KASUS & NOTIFIKASI\*\*  
\(HTTPS://LIBRARY.GUNADARMA.AC.ID/DEPC  
SYSTEM/ANGGOTA/NOTIFIKASI\)](https://library.gunadarma.ac.id/DEPC/system/anggota/notifikasi)

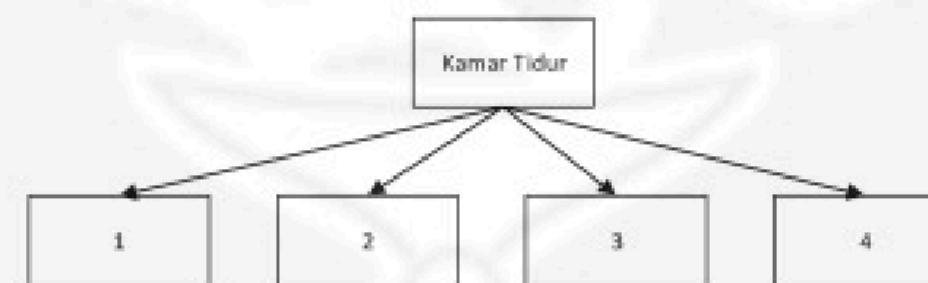
[\*\*BEBAS PERPUSTAKAAN\*\*  
\(HTTPS://LIBRARY.GUNADARMA.AC.ID/DEPC  
SYSTEM/ANGGOTA/BEBASPERPUSTAKAAN\)](https://library.gunadarma.ac.id/DEPC/system/anggota/bebasperpustakaan)

[\*\*KUNJUNGAN\*\*  
\(HTTPS://LIBRARY.GUNADARMA.AC.ID/DEPC  
SYSTEM/KUNJUNGAN\)](https://library.gunadarma.ac.id/DEPC/system/kunjungan)

[\*\*E-PAPER\*\*  
\(HTTPS://LIBRARY.GUNADARMA.AC.ID/DEPC\)](https://library.gunadarma.ac.id/DEPC)

4	9E+16	2	0.133333333	1800000000
weight avg mse	8.37921E+16			
<b>KAMAR MANDI</b>	<b>MSE (HARGA)</b>	<b>COUNT</b>	<b>PROP</b>	<b>AVG</b>
1	8.66542E+16	5	0.333333333	566600000
2	2.55625E+17	6	0.4	723000000
3	1.3095E+17	4	0.266666667	1680000000
weight avg mse	1.66055E+17			
<b>GARASI</b>	<b>MSE (HARGA)</b>	<b>COUNT</b>	<b>PROP</b>	<b>AVG</b>
0	4.27361E+17	6	0.4	926333333.3
1	2.40753E+17	6	0.4	928000000
2	5.49739E+17	3	0.2	921666666.7
weight avg mse	3.77193E+17			
<b>HASIL SPLITTING</b>	<b>8.37921E+16</b>			
<b>MSE TERKECIL ADALAH KAMAR TIDUR</b>				

Berdasarkan Tabel 3.15 di atas hasil MSE terkecil di dapatkan pada kolom kamar tidur sehingga pada kolom ini menjadi variabel yang akan melakukan percabangan atau terpilih menjadi *splitting attribute*. Hasil percabangan ini akan dibagi berdasarkan jumlah kamar tidur.



Gambar 3. 13 Hasil *Splitting* Variabel Kamar Tidur

Berdasarkan Gambar 3.13 di atas adalah hasil dari percabangan variabel kamar tidur menjadi 4 sesuai dengan jumlahnya masing-masing. Langkah selanjutnya akan dilakukan pembagian data berdasarkan kategori jumlah kamar tidur dikarenakan simpul yang terbentuk belum memenuhi *splitting criteria* dan harus melalui proses *splitting* hingga dalam suatu simpul hanya terdapat satu sampel.

Tabel 3. 16 Data Berdasarkan Jumlah Kamar Tidur

kamar tidur	kamar mandi	garasi	harga
1	2	2	230000000
1	2	0	168000000
2	1	1	268000000

HOME  
(HTTPS://LIBRARY.GUNADARMA.AC.ID/DEPC  
SYSTEM/)

PEMBAYARAN  
(HTTPS://LIBRARY.GUNADARMA.AC.ID/DEPC  
SYSTEM/PEMBAYARAN)

BIODATA DIRI  
(HTTPS://LIBRARY.GUNADARMA.AC.ID/DEPC  
SYSTEM/ANGGOTA)

PEMINJAMAN BUKU  
(HTTPS://LIBRARY.GUNADARMA.AC.ID/DEPC  
SYSTEM/ANGGOTA/PEMINJAMAN)

PENULISAN ILMIAH  
(HTTPS://LIBRARY.GUNADARMA.AC.ID/DEPC  
SYSTEM/ANGGOTA/PENULISAN)

FOTOCOPY  
(HTTPS://LIBRARY.GUNADARMA.AC.ID/DEPC  
SYSTEM/ANGGOTA/FOTOCOPY)

SUMBANG BUKU  
(HTTPS://LIBRARY.GUNADARMA.AC.ID/DEPC  
SYSTEM/ANGGOTA/SUMBANGBUKU)

KASUS & NOTIFIKASI  
(HTTPS://LIBRARY.GUNADARMA.AC.ID/DEPC  
SYSTEM/ANGGOTA/NOTIFIKASI)

BEBAS PERPUSTAKAAN  
(HTTPS://LIBRARY.GUNADARMA.AC.ID/DEPC  
SYSTEM/ANGGOTA/BEBASPERPUSTAKAAN)

KUNJUNGAN  
(HTTPS://LIBRARY.GUNADARMA.AC.ID/DEPC  
SYSTEM/KUNJUNGAN)

E-PAPER  
(HTTPS://LIBRARY.GUNADARMA.AC.ID/DEPC

2	1	0	2000000000
2	1	2	5850000000
2	2	1	3300000000
3	1	1	9000000000
3	3	0	12000000000
3	3	2	19500000000
3	3	1	14700000000
3	2	1	11000000000
3	2	0	10100000000
3	1	0	8800000000
4	2	1	15000000000
4	3	0	21000000000

Berdasarkan Tabel 3.16 di atas merupakan data *bootstrap* yang diurutkan berdasarkan jumlah kamar tidur. Jika sudah, maka ambil satu persatu sampel data berdasarkan jumlah kamar tidurnya dalam hal ini pertama ambil data dengan jumlah kamar tidur = 1.

Tabel 3. 17 Data Percabangan Variabel Kamar Tidur = 1

kamar tidur	kamar mandi	garasi	harga
1	2	2	2300000000
1	2	0	1680000000

Berdasarkan Tabel 3.17 simpul pertama dari variabel kamar tidur yang akan dilakukan percabangan adalah variabel kamar tidur = 1. Hitung nilai MSE terkecil dengan langkah perhitungan seperti sebelumnya menggunakan 2 variabel untuk melanjutkan simpul percabangan. Data dari masing-masing fitur akan dilakukan MSE untuk mendapatkan *splitting criteria* yang sesuai dengan pembentukan pohon *Random Forest*.

Tabel 3. 18 Hasil Perhitungan MSE Terkecil Percabangan  
Variabel Kamar Tidur = 1

KAMAR MANDI	MSE	COUNT	PROP	AVG
2	9.61E+14	2	1	1990000000
weight avg mse	9.61E+14			

HOME  
(HTTPS://LIBRARY.GUNADARMA.AC.ID/DEPC  
SYSTEM/)

PEMBAYARAN  
(HTTPS://LIBRARY.GUNADARMA.AC.ID/DEPC  
SYSTEM/PEMBAYARAN)

BIODATA DIRI  
(HTTPS://LIBRARY.GUNADARMA.AC.ID/DEPC  
SYSTEM/ANGGOTA)

PEMINJAMAN BUKU  
(HTTPS://LIBRARY.GUNADARMA.AC.ID/DEPC  
SYSTEM/ANGGOTA/PEMINJAMAN)

PENULISAN ILMIAH  
(HTTPS://LIBRARY.GUNADARMA.AC.ID/DEPC  
SYSTEM/ANGGOTA/PENULISAN)

FOTOCOPY  
(HTTPS://LIBRARY.GUNADARMA.AC.ID/DEPC  
SYSTEM/ANGGOTA/FOTOCOPY)

SUMBANG BUKU  
(HTTPS://LIBRARY.GUNADARMA.AC.ID/DEPC  
SYSTEM/ANGGOTA/SUMBANGBUKU)

KASUS & NOTIFIKASI  
(HTTPS://LIBRARY.GUNADARMA.AC.ID/DEPC  
SYSTEM/ANGGOTA/NOTIFIKASI)

BEBAS PERPUSTAKAAN  
(HTTPS://LIBRARY.GUNADARMA.AC.ID/DEPC  
SYSTEM/ANGGOTA/BEBASPERPUSTAKAAN)

KUNJUNGAN  
(HTTPS://LIBRARY.GUNADARMA.AC.ID/DEPC  
SYSTEM/KUNJUNGAN)

E-PAPER  
(HTTPS://LIBRARY.GUNADARMA.AC.ID/DEPC

GARASI	MSE	COUNT	PROP	AVG
0	0	1	0.5	160000000
2	0	1	0.5	230000000
weight avg mse	0			
<b>HASIL SPLITTING</b>				0
<b>VARIABEL GARASI MENJADI HASIL SPLIT 1</b>				

Berdasarkan Tabel 3.18 di atas hasil *splitting* dengan nilai *weight average* MSE terkecil didapatkan pada variabel jumlah garasi. Variabel jumlah garasi ini akan menjadi variabel percabangan selanjutnya dan akan dihitung nilai MSE terkecil untuk variabel kamar tidur = 1 untuk variabel garasi.

Tabel 3. 19 Data Percabangan Variabel Kamar Tidur = 1 untuk Variabel Garasi

GARASI	MSE	COUNT	KAMAR MANDI	AVG
0	0	1	2	168000000
2	0	1	2	230000000
WEIGHT	0			
<b>SUDAH MEMENUHI STOPING CRITERIA</b>				

Berdasarkan Tabel 3.19 di atas menunjukkan data dari variabel kamar tidur = 1 untuk variabel jumlah garasi sudah memenuhi *stopping criteria* terlihat dari masing-masing kategori yang akan dijadikan simpul memiliki satu sampel di setiap kategorinya. Pada jumlah garasi = 0 memiliki satu sampel dengan harga 168.000.000 dan untuk kategori jumlah garasi = 2 memiliki satu sampel dengan harga 230.000.000. Kesimpulan untuk rumah dengan jumlah kamar tidur = 1, jumlah garasi = 0 dan jumlah kamar mandi = 2 mendapatkan harga prediksi sebesar 168.000.000 dan untuk rumah dengan jumlah kamar tidur = 1, jumlah garasi = 2 dan jumlah kamar mandi = 2 mendapatkan harga prediksi sebesar 230.000.000.

[HOME  
\(HTTPS://LIBRARY.GUNADARMA.AC.ID/DEPC  
SYSTEM/\)](https://library.gunadarma.ac.id/DEPC/system/)

[PEMBAYARAN  
\(HTTPS://LIBRARY.GUNADARMA.AC.ID/DEPC  
SYSTEM/PEMBAYARAN\)](https://library.gunadarma.ac.id/DEPC/system/PEMBAYARAN)

[BIODATA DIRI  
\(HTTPS://LIBRARY.GUNADARMA.AC.ID/DEPC  
SYSTEM/ANGGOTA\)](https://library.gunadarma.ac.id/DEPC/system/ANGGOTA)

[PEMINJAMAN BUKU  
\(HTTPS://LIBRARY.GUNADARMA.AC.ID/DEPC  
SYSTEM/ANGGOTA/PEMINJAMAN\)](https://library.gunadarma.ac.id/DEPC/system/ANGGOTA/PEMINJAMAN)

[PENULISAN ILMIAH  
\(HTTPS://LIBRARY.GUNADARMA.AC.ID/DEPC  
SYSTEM/ANGGOTA/PENULISAN\)](https://library.gunadarma.ac.id/DEPC/system/ANGGOTA/PENULISAN)

[FOTOCOPY  
\(HTTPS://LIBRARY.GUNADARMA.AC.ID/DEPC  
SYSTEM/ANGGOTA/FOTOCOPY\)](https://library.gunadarma.ac.id/DEPC/system/ANGGOTA/FOTOCOPY)

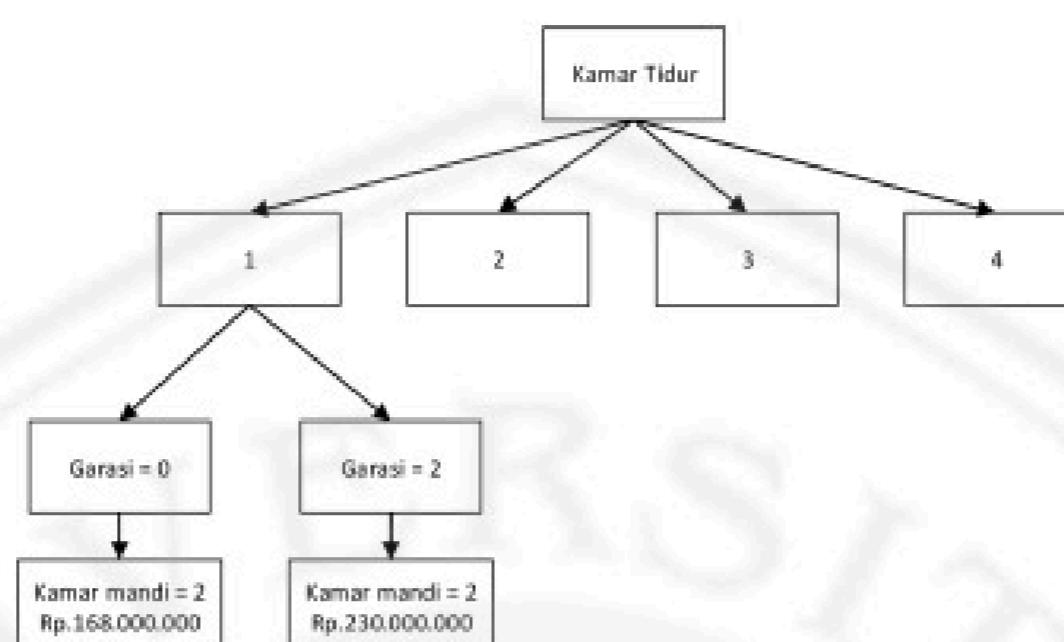
[SUMBANG BUKU  
\(HTTPS://LIBRARY.GUNADARMA.AC.ID/DEPC  
SYSTEM/ANGGOTA/SUMBANGBUKU\)](https://library.gunadarma.ac.id/DEPC/system/ANGGOTA/SUMBANGBUKU)

[KASUS & NOTIFIKASI  
\(HTTPS://LIBRARY.GUNADARMA.AC.ID/DEPC  
SYSTEM/ANGGOTA/NOTIFIKASI\)](https://library.gunadarma.ac.id/DEPC/system/ANGGOTA/NOTIFIKASI)

[BEBAS PERPUSTAKAAN  
\(HTTPS://LIBRARY.GUNADARMA.AC.ID/DEPC  
SYSTEM/ANGGOTA/BEBASPERPUSTAKAAN\)](https://library.gunadarma.ac.id/DEPC/system/ANGGOTA/BEBASPERPUSTAKAAN)

[KUNJUNGAN  
\(HTTPS://LIBRARY.GUNADARMA.AC.ID/DEPC  
SYSTEM/KUNJUNGAN\)](https://library.gunadarma.ac.id/DEPC/system/KUNJUNGAN)

[E-PAPER  
\(HTTPS://LIBRARY.GUNADARMA.AC.ID/DEPC\)](https://library.gunadarma.ac.id/DEPC)



Gambar 3. 14 Hasil Percabangan Variabel Kamar Tidur = 1 untuk Variabel Garasi

Berdasarkan Gambar 3.14 di atas dapat dilihat bahwa variabel jumlah kamar tidur = 1 sudah mencapai terminal node. Langkah selanjutnya akan dilakukan perhitungan terhadap variabel jumlah kamar tidur = 2 dengan cara yang sama seperti sebelumnya.

Tabel 3. 20 Data Percabangan Variabel Kamar Tidur = 2

kamar tidur	kamar mandi	garasi	harga
2	1	1	268000000
2	1	0	200000000
2	1	2	585000000
2	2	1	330000000

Berdasarkan Tabel 3.20 di atas merupakan data percabangan untuk variabel kamar tidur = 2. Lakukan pencarian nilai MSE terkecil menggunakan cara yang sama seperti sebelumnya pada setiap data kamar mandi dan garasi agar mendapatkan hasil *splitting* data untuk pembentukan pohon selanjutnya.

Tabel 3. 21 Menghitung Nilai MSE Terkecil Percabangan

Variabel Kamar Mandi = 2

KAMAR MANDI	MSE	COUNT	PROP	AVG
1	2.81487E+16	3	0.75	351000000
2	0	1	0.25	330000000

[HOME  
\(HTTPS://LIBRARY.GUNADARMA.AC.ID/DEPC/SYSTEM/\)](https://library.gunadarma.ac.id/DEPC/SYSTEM/)

[PEMBAYARAN  
\(HTTPS://LIBRARY.GUNADARMA.AC.ID/DEPC/SYSTEM/PEMBAYARAN\)](https://library.gunadarma.ac.id/DEPC/SYSTEM/PEMBAYARAN)

[BIODATA DIRI  
\(HTTPS://LIBRARY.GUNADARMA.AC.ID/DEPC/SYSTEM/ANGGOTA\)](https://library.gunadarma.ac.id/DEPC/SYSTEM/ANGGOTA)

[PEMINJAMAN BUKU  
\(HTTPS://LIBRARY.GUNADARMA.AC.ID/DEPC/SYSTEM/ANGGOTA/PEMINJAMAN\)](https://library.gunadarma.ac.id/DEPC/SYSTEM/ANGGOTA/PEMINJAMAN)

[PENULISAN ILMIAH  
\(HTTPS://LIBRARY.GUNADARMA.AC.ID/DEPC/SYSTEM/ANGGOTA/PENULISAN\)](https://library.gunadarma.ac.id/DEPC/SYSTEM/ANGGOTA/PENULISAN)

[FOTOCOPY  
\(HTTPS://LIBRARY.GUNADARMA.AC.ID/DEPC/SYSTEM/ANGGOTA/FOTOCOPY\)](https://library.gunadarma.ac.id/DEPC/SYSTEM/ANGGOTA/FOTOCOPY)

[SUMBANG BUKU  
\(HTTPS://LIBRARY.GUNADARMA.AC.ID/DEPC/SYSTEM/ANGGOTA/SUMBANGBUKU\)](https://library.gunadarma.ac.id/DEPC/SYSTEM/ANGGOTA/SUMBANGBUKU)

[KASUS & NOTIFIKASI  
\(HTTPS://LIBRARY.GUNADARMA.AC.ID/DEPC/SYSTEM/ANGGOTA/NOTIFIKASI\)](https://library.gunadarma.ac.id/DEPC/SYSTEM/ANGGOTA/NOTIFIKASI)

[BEBAS PERPUSTAKAAN  
\(HTTPS://LIBRARY.GUNADARMA.AC.ID/DEPC/SYSTEM/ANGGOTA/BEBASPERPUSTAKAAN\)](https://library.gunadarma.ac.id/DEPC/SYSTEM/ANGGOTA/BEBASPERPUSTAKAAN)

[KUNJUNGAN  
\(HTTPS://LIBRARY.GUNADARMA.AC.ID/DEPC/SYSTEM/KUNJUNGAN\)](https://library.gunadarma.ac.id/DEPC/SYSTEM/KUNJUNGAN)

[E-PAPER  
\(HTTPS://LIBRARY.GUNADARMA.AC.ID/DEPC/E-PAPER\)](https://library.gunadarma.ac.id/DEPC/E-PAPER)

WEIGHT	2.11115E+16	COUNT	PROP	AVG
GARASI	MSE	COUNT	PROP	AVG
0	0	1	0.25	200000000
1	1.00494E+16	2	0.50	394333333.3
2	0	1	0.25	585000000
WEIGHT	5.02472E+15			
HASIL SPLITTING 2	5.02472E+15			
VARIABEL GARASI MENJADI HASIL SPLIT 2				

Berdasarkan Tabel 3.21 di atas hasil mencari nilai MSE terkecil didapatkan oleh variabel garasi maka akan dilakukan perhitungan dalam mencari *stopping criteria* dari variabel kamar mandi = 2 untuk variabel garasi.

Tabel 3. 22 Data Percabangan Variabel Kamar Tidur = 2 untuk Variabel Garasi

GARASI	MSE	COUNT	KAMAR MANDI	AVG
0	0	1	1	200000000
1	1.9105E+16	2	1	394333333.3
2	0	1	2	585000000
<b>GARASI 0 DAN 2 SUDAH MEMENUHI CRITERIA</b>				

Berdasarkan Tabel 3.22 di atas untuk variabel garasi = 0 dan 2 sudah memenuhi *stopping criteria* sedangkan untuk variabel garasi = 1 harus dilakukan pencarian nilai MSE terkecil lagi agar mencapai kondisi *stopping criteria*. Perhitungan variabel garasi = 1 akan menggunakan fitur kamar mandi sebagai pendukung untuk memenuhi *criteria*.

Tabel 3. 23 Data Percabangan Variabel Kamar Tidur = 2 untuk Variabel Garasi dan Variabel Kamar Mandi

KAMAR MANDI	MSE	COUNT	GARASI	AVG
1	0	1	1	268000000
2	0	1	1	330000000
<b>DATA GARASI 1 DENGAN KAMAR MANDI 1 SUDAH MEMENUHI CRITERIA</b>				
<b>DATA GARASI 1 DENGAN KAMAR MANDI 2 SUDAH MEMENUHI CRITERIA</b>				

HOME  
(HTTP://LIBRARY.GUNADARMA.AC.ID/DEPC SYSTEM/)

PEMBAYARAN  
(HTTP://LIBRARY.GUNADARMA.AC.ID/DEPC SYSTEM/PEMBAYARAN)

BIODATA DIRI  
(HTTP://LIBRARY.GUNADARMA.AC.ID/DEPC SYSTEM/ANGGOTA)

PEMINJAMAN BUKU  
(HTTP://LIBRARY.GUNADARMA.AC.ID/DEPC SYSTEM/ANGGOTA/PEMINJAMAN)

PENULISAN ILMIAH  
(HTTP://LIBRARY.GUNADARMA.AC.ID/DEPC SYSTEM/ANGGOTA/PENULISAN)

FOTOCOPY  
(HTTP://LIBRARY.GUNADARMA.AC.ID/DEPC SYSTEM/ANGGOTA/FOTOCOPY)

SUMBANG BUKU  
(HTTP://LIBRARY.GUNADARMA.AC.ID/DEPC SYSTEM/ANGGOTA/SUMBANGBUKU)

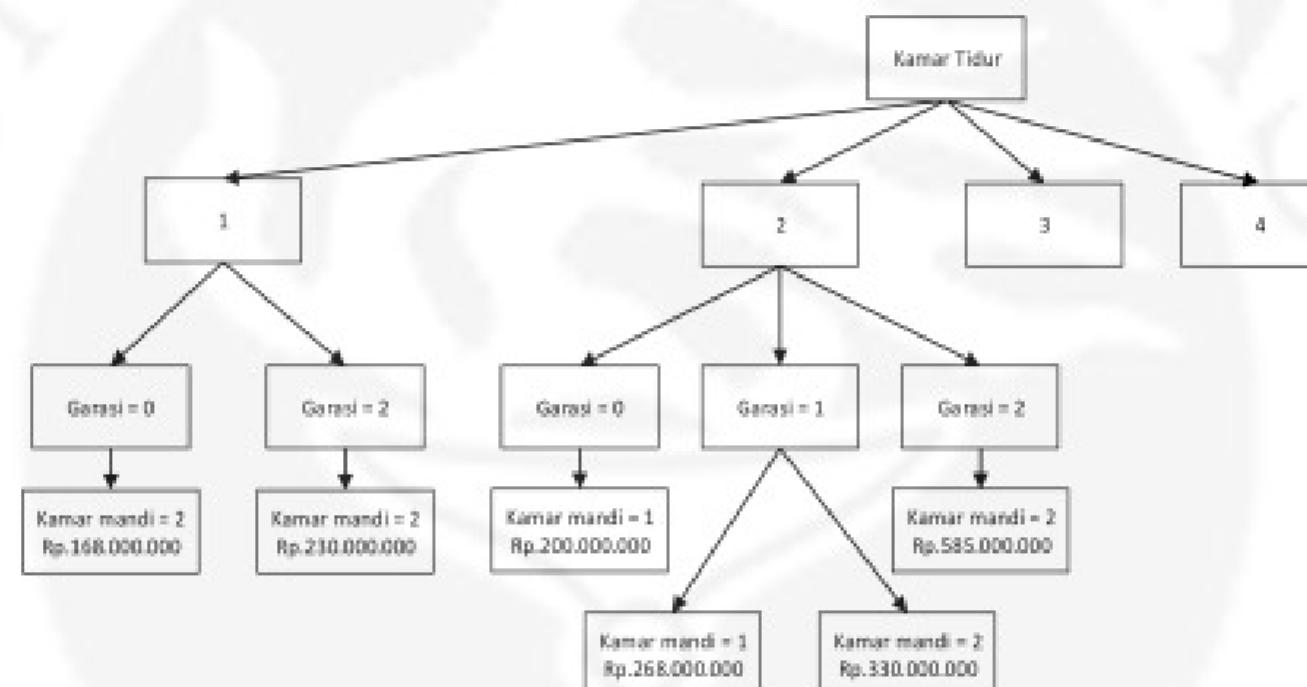
KASUS & NOTIFIKASI  
(HTTP://LIBRARY.GUNADARMA.AC.ID/DEPC SYSTEM/ANGGOTA/NOTIFIKASI)

BEBAS PERPUSTAKAAN  
(HTTP://LIBRARY.GUNADARMA.AC.ID/DEPC SYSTEM/ANGGOTA/BEBASTERPUSTAKAAN)

KUNJUNGAN  
(HTTP://LIBRARY.GUNADARMA.AC.ID/DEPC SYSTEM/KUNJUNGAN)

E-PAPER  
(HTTP://LIBRARY.GUNADARMA.AC.ID/DEPC

Berdasarkan Tabel 3.23 di atas dapat dilihat bahwa untuk variabel garasi = 1 sudah mencapai *stopping criteria* dengan memiliki satu sampel pada masing-masing data. Disimpulkan harga rumah untuk jumlah kamar tidur = 2, jumlah garasi = 0 dan jumlah kamar mandi = 1 di prediksi sebesar 200.000.000. Harga rumah dengan jumlah kamar tidur = 2, jumlah garasi = 2 dan jumlah kamar mandi = 2 di prediksi sebesar 585.000.000. Sedangkan untuk harga rumah dengan jumlah kamar mandi = 2, jumlah garasi = 1 dan jumlah kamar mandi = 1 di prediksi sebesar 268.000.000. Harga rumah dengan jumlah kamar mandi = 2, jumlah garasi = 1 dan jumlah kamar mandi 330.000.000.



Gambar 3. 15 Hasil Percabangan Variabel Kamar Tidur = 2 untuk Variabel Garasi

Berdasarkan Gambar 3.15 di atas merupakan hasil percabangan dengan variabel kamar tidur = 2 masing-masing sudah mencapai terminal *node*. Setiap daerah percabangan memiliki *stopping criteria* masing-masing. Lakukan perhitungan terhadap variabel kamar tidur = 3 dan variabel kamar tidur = 4 untuk melengkapi *decision tree* yang terbentuk.

[HOME  
\(HTTPS://LIBRARY.GUNADARMA.AC.ID/DEPC  
SYSTEM/\)](https://library.gunadarma.ac.id/DEPC/system/)

[PEMBAYARAN  
\(HTTPS://LIBRARY.GUNADARMA.AC.ID/DEPC  
SYSTEM/PEMBAYARAN\)](https://library.gunadarma.ac.id/DEPC/system/PEMBAYARAN)

[BIODATA DIRI  
\(HTTPS://LIBRARY.GUNADARMA.AC.ID/DEPC  
SYSTEM/ANGGOTA\)](https://library.gunadarma.ac.id/DEPC/system/ANGGOTA)

[PEMINJAMAN BUKU  
\(HTTPS://LIBRARY.GUNADARMA.AC.ID/DEPC  
SYSTEM/ANGGOTA/PEMINJAMAN\)](https://library.gunadarma.ac.id/DEPC/system/ANGGOTA/PEMINJAMAN)

[PENULISAN ILMIAH  
\(HTTPS://LIBRARY.GUNADARMA.AC.ID/DEPC  
SYSTEM/ANGGOTA/PENULISAN\)](https://library.gunadarma.ac.id/DEPC/system/ANGGOTA/PENULISAN)

[FOTOCOPY  
\(HTTPS://LIBRARY.GUNADARMA.AC.ID/DEPC  
SYSTEM/ANGGOTA/FOTOCOPY\)](https://library.gunadarma.ac.id/DEPC/system/ANGGOTA/FOTOCOPY)

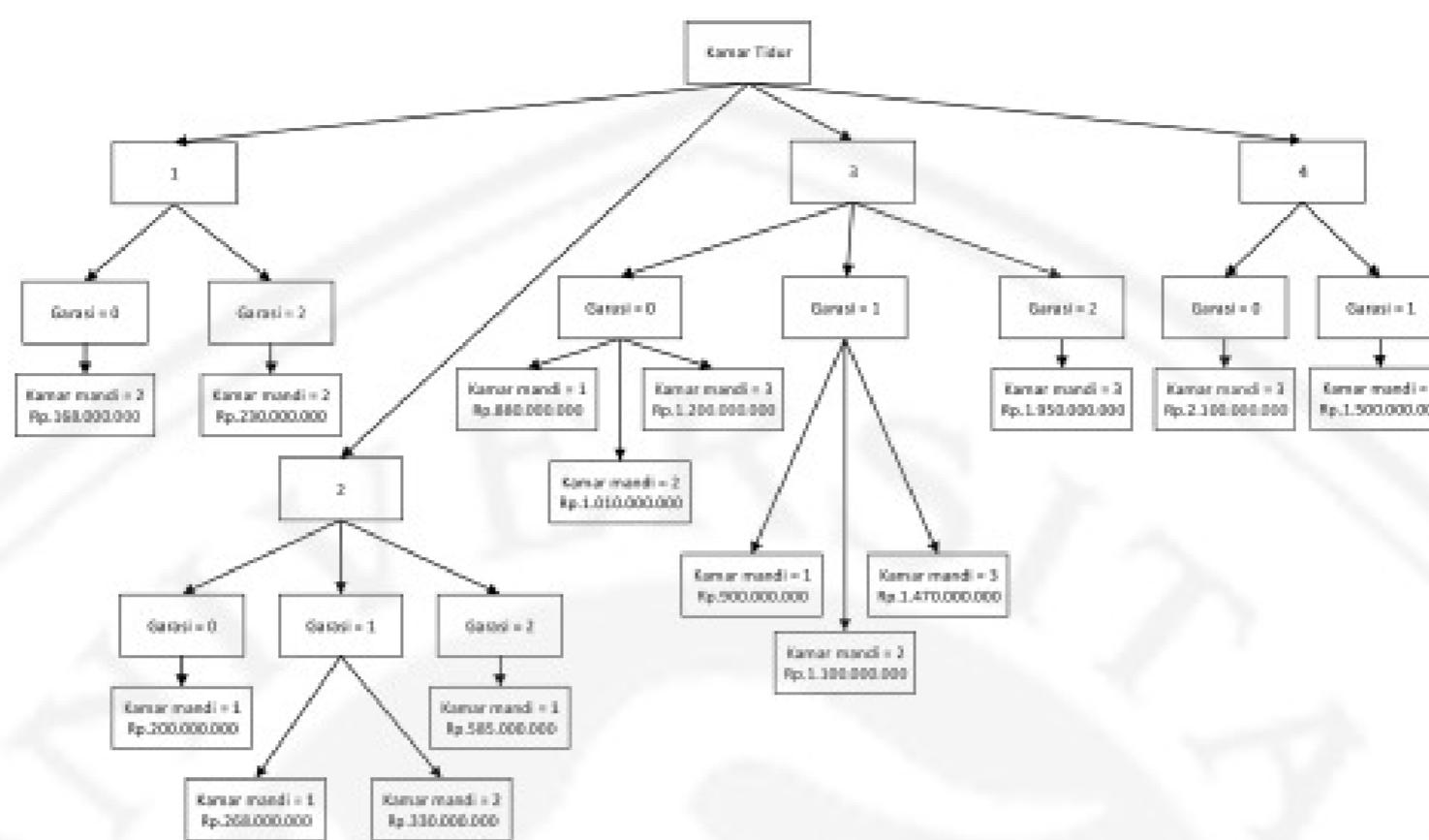
[SUMBANG BUKU  
\(HTTPS://LIBRARY.GUNADARMA.AC.ID/DEPC  
SYSTEM/ANGGOTA/SUMBANGBUKU\)](https://library.gunadarma.ac.id/DEPC/system/ANGGOTA/SUMBANGBUKU)

[KASUS & NOTIFIKASI  
\(HTTPS://LIBRARY.GUNADARMA.AC.ID/DEPC  
SYSTEM/ANGGOTA/NOTIFIKASI\)](https://library.gunadarma.ac.id/DEPC/system/ANGGOTA/NOTIFIKASI)

[BEBAS PERPUSTAKAAN  
\(HTTPS://LIBRARY.GUNADARMA.AC.ID/DEPC  
SYSTEM/ANGGOTA/BEBASPERPUSTAKAAN\)](https://library.gunadarma.ac.id/DEPC/system/ANGGOTA/BEBASPERPUSTAKAAN)

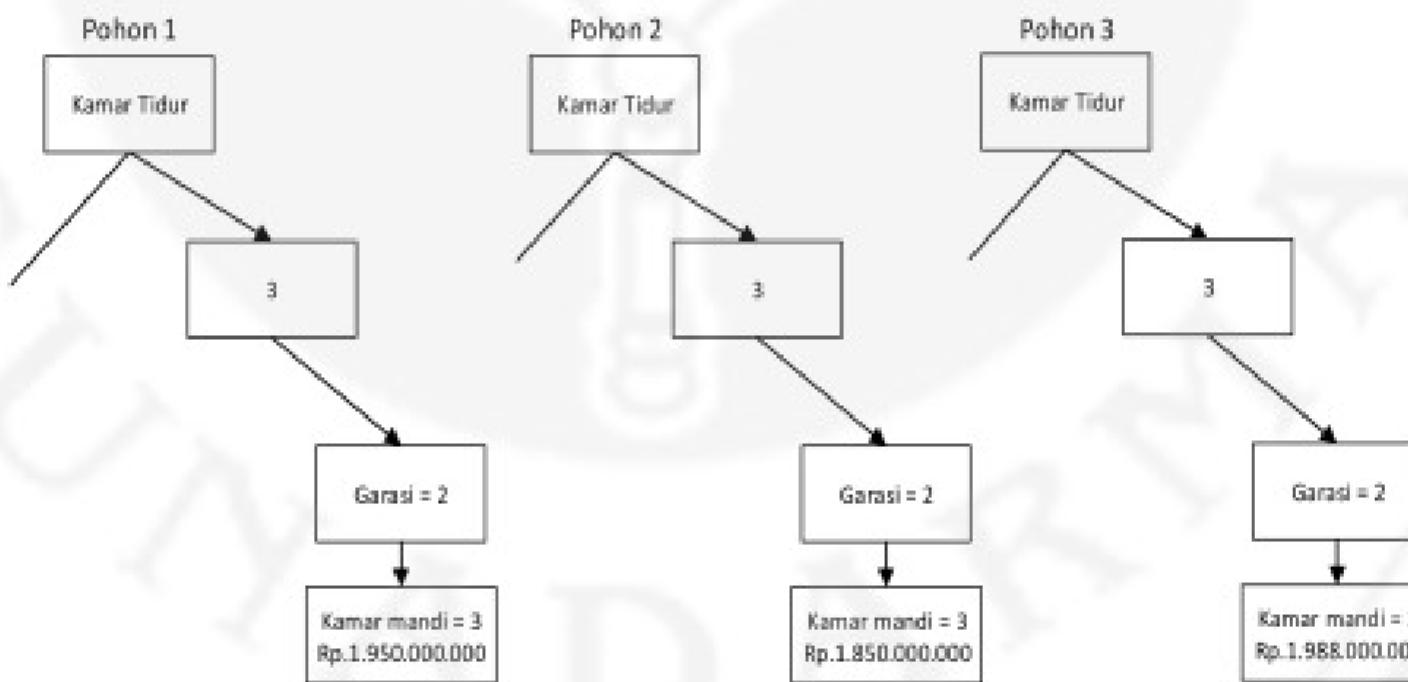
[KUNJUNGAN  
\(HTTPS://LIBRARY.GUNADARMA.AC.ID/DEPC  
SYSTEM/KUNJUNGAN\)](https://library.gunadarma.ac.id/DEPC/system/KUNJUNGAN)

[E-PAPER  
\(HTTPS://LIBRARY.GUNADARMA.AC.ID/DEPC\)](https://library.gunadarma.ac.id/DEPC)



Gambar 3. 16 Hasil *Decision Tree* Perhitungan Manual Prediksi Harga Rumah

Berdasarkan Gambar 3.16 di atas adalah hasil dari perhitungan keseluruhan pada 15 data *bootstrap* sampel yang digunakan untuk perhitungan manual. *Random Forest* merupakan hasil *aggregate* dari banyak pohon *decision tree* maka dari itu dibuatlah beberapa pohon *decision tree* berdasarkan 3 kumpulan data sampel berbeda dengan jumlah yang sama.



Gambar 3. 17 Pohon Aggregate Prediksi *Random Forest*

Berdasarkan Gambar 3.17 merupakan kumpulan dari beberapa pohon yang memiliki cabang *stopping criteria* yang sama kemudian akan dicari rata-rata berdasarkan Persamaan 2.2.

[HOME  
\(HTTPS://LIBRARY.GUNADARMA.AC.ID/DEPC  
SYSTEM/\)](https://library.gunadarma.ac.id/DEPC/system/)

[PEMBAYARAN  
\(HTTPS://LIBRARY.GUNADARMA.AC.ID/DEPC  
SYSTEM/PEMBAYARAN\)](https://library.gunadarma.ac.id/DEPC/system/PEMBAYARAN)

[BIODATA DIRI  
\(HTTPS://LIBRARY.GUNADARMA.AC.ID/DEPC  
SYSTEM/ANGGOTA\)](https://library.gunadarma.ac.id/DEPC/system/ANGGOTA)

[PEMINJAMAN BUKU  
\(HTTPS://LIBRARY.GUNADARMA.AC.ID/DEPC  
SYSTEM/ANGGOTA/PEMINJAMAN\)](https://library.gunadarma.ac.id/DEPC/system/ANGGOTA/PEMINJAMAN)

[PENULISAN ILMIAH  
\(HTTPS://LIBRARY.GUNADARMA.AC.ID/DEPC  
SYSTEM/ANGGOTA/PENULISAN\)](https://library.gunadarma.ac.id/DEPC/system/ANGGOTA/PENULISAN)

[FOTOCOPY  
\(HTTPS://LIBRARY.GUNADARMA.AC.ID/DEPC  
SYSTEM/ANGGOTA/FOTOCOPY\)](https://library.gunadarma.ac.id/DEPC/system/ANGGOTA/FOTOCOPY)

[SUMBANG BUKU  
\(HTTPS://LIBRARY.GUNADARMA.AC.ID/DEPC  
SYSTEM/ANGGOTA/SUMBANGBUKU\)](https://library.gunadarma.ac.id/DEPC/system/ANGGOTA/SUMBANGBUKU)

[KASUS & NOTIFIKASI  
\(HTTPS://LIBRARY.GUNADARMA.AC.ID/DEPC  
SYSTEM/ANGGOTA/NOTIFIKASI\)](https://library.gunadarma.ac.id/DEPC/system/ANGGOTA/NOTIFIKASI)

[BEBAS PERPUSTAKAAN  
\(HTTPS://LIBRARY.GUNADARMA.AC.ID/DEPC  
SYSTEM/ANGGOTA/BEBASPERPUSTAKAAN\)](https://library.gunadarma.ac.id/DEPC/system/ANGGOTA/BEBASPERPUSTAKAAN)

[KUNJUNGAN  
\(HTTPS://LIBRARY.GUNADARMA.AC.ID/DEPC  
SYSTEM/KUNJUNGAN\)](https://library.gunadarma.ac.id/DEPC/system/KUNJUNGAN)

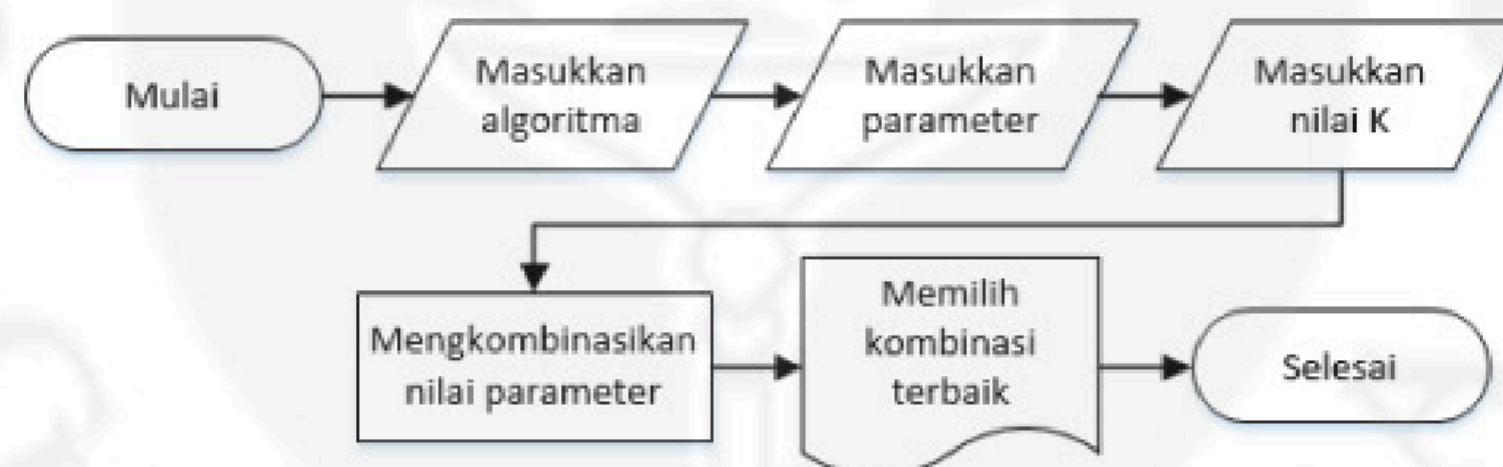
[E-PAPER  
\(HTTPS://LIBRARY.GUNADARMA.AC.ID/DEPC\)](https://library.gunadarma.ac.id/DEPC)

$$\begin{aligned}\hat{Y}_i &= \frac{1}{N_{tree}} \sum_{n=1}^{N_{tree}} \hat{Y}_n \\ \hat{Y}_i &= \frac{1}{3} (1.950.000.000 + 1.850.000.000 + 1.988.000.000) \\ &= \frac{5.788.000.000}{3} = 1.929.333.333\end{aligned}$$

Hasil akhir dari prediksi *Random Forest* pada Harga Rumah dengan jumlah kamar tidur = 3, jumlah garasi = 2, dan jumlah kamar mandi = 3 sebesar 1.929.333.333.

### 3.4.3. Tuning parameter

Pemodelan *Random Forest Regression* yang telah dilakukan menggunakan parameter yang ditentukan secara manual oleh peneliti, untuk menguji *performance* dari parameter model yang sudah dibuat maka dari itu dibutuhkan *tuning parameter*. Pada kasus ini akan menggunakan *tuning parameter* dengan metode *searchGridCV* dan matriks yang digunakan untuk pengukuran performa dari *Grid Search* adalah *Mean Cross Validation*.



Gambar 3. 18 Proses *Grid Search CV*

Berdasarkan Gambar 3.18 di atas proses yang akan dilakukan oleh metode *grid cv* ini pertama-tama menentukan algoritma yang akan digunakan untuk proses *tuning parameter*, kemudian memasukkan parameter yang disesuaikan oleh algoritma yang digunakan. Langkah selanjutnya adalah masukkan jumlah nilai *k-fold validation* yang akan di uji coba. Jumlah nilai *k* ini akan menentukan pembentukan ukuran data yang sama pada setiap *fold*. Data pada setiap *fold* akan dikombinasikan dengan parameter yang digunakan dan *grid search* akan memilih kombinasi terbaik berdasarkan nilai *CV Score* yang tertinggi.

HOME  
(HTTPS://LIBRARY.GUNADARMA.AC.ID/DEPC  
SYSTEM/)

PEMBAYARAN  
(HTTPS://LIBRARY.GUNADARMA.AC.ID/DEPC  
SYSTEM/PEMBAYARAN)

BIODATA DIRI  
(HTTPS://LIBRARY.GUNADARMA.AC.ID/DEPC  
SYSTEM/ANGGOTA)

PEMINJAMAN BUKU  
(HTTPS://LIBRARY.GUNADARMA.AC.ID/DEPC  
SYSTEM/ANGGOTA/PEMINJAMAN)

PENULISAN ILMIAH  
(HTTPS://LIBRARY.GUNADARMA.AC.ID/DEPC  
SYSTEM/ANGGOTA/PENULISAN)

FOTOCOPY  
(HTTPS://LIBRARY.GUNADARMA.AC.ID/DEPC  
SYSTEM/ANGGOTA/FOTOCOPY)

SUMBANG BUKU  
(HTTPS://LIBRARY.GUNADARMA.AC.ID/DEPC  
SYSTEM/ANGGOTA/SUMBANGBUKU)

KASUS & NOTIFIKASI  
(HTTPS://LIBRARY.GUNADARMA.AC.ID/DEPC  
SYSTEM/ANGGOTA/NOTIFIKASI)

BEBAS PERPUSTAKAAN  
(HTTPS://LIBRARY.GUNADARMA.AC.ID/DEPC  
SYSTEM/ANGGOTA/BEBASTERPUSTAKAAN)

KUNJUNGAN  
(HTTPS://LIBRARY.GUNADARMA.AC.ID/DEPC  
SYSTEM/KUNJUNGAN)

E-PAPER  
(HTTPS://LIBRARY.GUNADARMA.AC.ID/DEPC

Beberapa parameter yang dibutuhkan pada algoritma *Random Forest Regression* sebagai berikut:

Tabel 3. 24 Parameter *Random Forest Regression*

Parameter	Pilihan
bootstrap	['True', 'False'],
criterion	['squared_error', 'absolute_error', 'poisson'],
max_features	[2, 3, 'auto', 'sqrt']
n_estimators	[100, 200, 300, 1000]
max_depth	[100, 200, 300, 400]
min_samples_split	[2, 6, 10]
min_samples_leaf	[1, 2, 3, 4]
random_state	[3, 12, 42]

Berdasarkan Tabel 3.24 di atas merupakan beberapa parameter penting yang harus ditentukan dalam pemodelan algoritma *Random Forest Regression*. Pada parameter *bootstrap* jika bernilai *true* maka akan menggunakan data sampel hasil pilihan *bootstrap*, jika *false* maka akan menggunakan semua dataset pada proses membangun pohon. Parameter *criterion* akan digunakan untuk mengukur kualitas dari *splitting* model dengan menggunakan beberapa *error*. Parameter *max\_features* digunakan untuk mempertimbangkan jumlah fitur pada saat mencari split terbaik. Pada kasus klasifikasi direkomendasikan untuk menggunakan *max\_features = sqrt(n\_features)*, sedangkan kasus regresi direkomendasikan untuk menggunakan *max\_features = n\_features/3*. Pada parameter selanjutnya menentukan jumlah pohon yang akan terbentuk menggunakan *n\_estimator* dengan nilai yang beragam. Parameter *max\_depth* digunakan untuk memberikan nilai maksimum level dari *decision tree* yang akan terbuat. Pada parameter *min\_samples\_split* digunakan untuk menentukan jumlah minimum split pohon, sedangkan *min\_samples\_leaf* untuk nilai minimum dari daun pada setiap pohon yang terbentuk. Parameter *random\_state* digunakan untuk menghasilkan data acak secara konsisten, apabila

HOME  
(HTTPS://LIBRARY.GUNADARMA.AC.ID/DEPC  
SYSTEM/)

PEMBAYARAN  
(HTTPS://LIBRARY.GUNADARMA.AC.ID/DEPC  
SYSTEM/PEMBAYARAN)

BIODATA DIRI  
(HTTPS://LIBRARY.GUNADARMA.AC.ID/DEPC  
SYSTEM/ANGGOTA)

PEMINJAMAN BUKU  
(HTTPS://LIBRARY.GUNADARMA.AC.ID/DEPC  
SYSTEM/ANGGOTA/PEMINJAMAN)

PENULISAN ILMIAH  
(HTTPS://LIBRARY.GUNADARMA.AC.ID/DEPC  
SYSTEM/ANGGOTA/PENULISAN)

FOTOCOPY  
(HTTPS://LIBRARY.GUNADARMA.AC.ID/DEPC  
SYSTEM/ANGGOTA/FOTOCOPY)

SUMBANG BUKU  
(HTTPS://LIBRARY.GUNADARMA.AC.ID/DEPC  
SYSTEM/ANGGOTA/SUMBANGBUKU)

KASUS & NOTIFIKASI  
(HTTPS://LIBRARY.GUNADARMA.AC.ID/DEPC  
SYSTEM/ANGGOTA/NOTIFIKASI)

BEBAS PERPUSTAKAAN  
(HTTPS://LIBRARY.GUNADARMA.AC.ID/DEPC  
SYSTEM/ANGGOTA/BEBASTERPUSTAKAAN)

KUNJUNGAN  
(HTTPS://LIBRARY.GUNADARMA.AC.ID/DEPC  
SYSTEM/KUNJUNGAN)

E-PAPER  
(HTTPS://LIBRARY.GUNADARMA.AC.ID/DEPC

parameter ini tidak di atur maka data latih yang akan di model menggunakan data acak yang berubah jika di jalankan kembali.

### 3.4.4. Modeling Algoritma Cosine Similarity

Algoritma *Cosine Similarity* ini akan digunakan untuk melihat hasil rekomendasi berdasarkan data inputan dari *user*. Pada proses perhitungan manual algoritma *Cosine Similarity* ini akan menggunakan 7 data sampel dan 1 data baru. Dimana data baru ini sudah mendapatkan hasil prediksi harga sehingga rekomendasi yang akan diberikan adalah berdasarkan beberapa faktor yang mendekati seperti luas tanah, luas bangunan, dan lainnya termasuk faktor harga akan mempengaruhi hasil prediksi.

Tabel 3. 25 Data Sampel Perhitungan *Cosine Similarity*

	Rumah minimalis di bandung	Rumah di Bandung	Cluster Mewah	Rumah Mezanin di Bandung	Rumah Cluster Kota Bekasi	Rumah Bandung Mewah	pengguna
daerah	0	0	8	0	8	0	0
harga	180000000 0	632000000 0	265000000 0	597000000 0	950000000 0	720000000 0	600000000 0
kamar tidur	2	2	2	2	3	2	2
kamar mandi	1	1	1	1	3	1	1
garasi	1	0	1	1	1	2	1
luas tanah	36	70	55	50	150	70	60
luas bangunan	50	50	36	40	90	60	56

Berdasarkan Tabel 3.25 di atas merupakan data sampel acak dan 1 data pengguna yang sudah mendapatkan hasil prediksi harga. Nilai *Cosine Similarity* pertama harus melakukan perkalian antara *query* (data baru) dan *item* (data sampel). (rumah minimalis di Bandung, pengguna)

$$\text{Kota} = 0 \times 0 = 0$$

$$\text{Harga} = 180.000.000 \times 600.000.000 = 1.08e + 17$$

$$\text{Kamar tidur} = 2 \times 2 = 4$$

$$\text{Kamar mandi} = 1 \times 1 = 1$$

$$\text{Garasi} = 1 \times 1 = 1$$

$$\text{Luas tanah} = 36 \times 60 = 2160$$

HOME  
(HTTPS://LIBRARY.GUNADARMA.AC.ID/DEPC  
SYSTEM/)

PEMBAYARAN  
(HTTPS://LIBRARY.GUNADARMA.AC.ID/DEPC  
SYSTEM/PEMBAYARAN)

BIODATA DIRI  
(HTTPS://LIBRARY.GUNADARMA.AC.ID/DEPC  
SYSTEM/ANGGOTA)

PEMINJAMAN BUKU  
(HTTPS://LIBRARY.GUNADARMA.AC.ID/DEPC  
SYSTEM/ANGGOTA/PEMINJAMAN)

PENULISAN ILMIAH  
(HTTPS://LIBRARY.GUNADARMA.AC.ID/DEPC  
SYSTEM/ANGGOTA/PENULISAN)

FOTOCOPY  
(HTTPS://LIBRARY.GUNADARMA.AC.ID/DEPC  
SYSTEM/ANGGOTA/FOTOCOPY)

SUMBANG BUKU  
(HTTPS://LIBRARY.GUNADARMA.AC.ID/DEPC  
SYSTEM/ANGGOTA/SUMBANGBUKU)

KASUS & NOTIFIKASI  
(HTTPS://LIBRARY.GUNADARMA.AC.ID/DEPC  
SYSTEM/ANGGOTA/NOTIFIKASI)

BEBAS PERPUSTAKAAN  
(HTTPS://LIBRARY.GUNADARMA.AC.ID/DEPC  
SYSTEM/ANGGOTA/BEBAZPERPUSTAKAAN)

KUNJUNGAN  
(HTTPS://LIBRARY.GUNADARMA.AC.ID/DEPC  
SYSTEM/KUNJUNGAN)

E-PAPER  
(HTTPS://LIBRARY.GUNADARMA.AC.ID/DEPC

$$\text{Luas bangunan} = 50 \times 56 = 2800$$

Jika semua atribut pada data sudah dikalikan maka jumlahkan semuanya. Lakukanlah perhitungan tersebut pada semua data sampel yang digunakan untuk mendapatkan hasil perhitungan dari pasangan kolom antara *query* dengan masing-masing *item*.

Tabel 3. 26 Hasil Perhitungan *Query\*Item*

	data(1,7)	data(2,7)	data(3,7)	data(4,7)	data(5,7)	data(6,7)
daerah	0	0	0	0	0	0
harga	1.08E+17	3.792E+17	1.59E+17	3.582E+17	5.7E+17	4.32E+17
kamar tidur	4	4	4	4	6	4
kamar mandi	1	1	1	1	3	1
garasi	1	0	1	1	1	2
luas tanah	2160	4200	3300	3000	9000	4200
luas bangunan	2800	2800	2016	2240	5040	3360
total	4966	7005	5322	5246	14050	7567

Berdasarkan Tabel 3.26 di atas adalah hasil perhitungan semua data sampel terhadap data baru. Langkah selanjutnya adalah menghitung kuadrat dari masing-masing *item* dan *query*.

Data (1):

$$\text{Kota} = 0^2 = 0$$

$$\text{Harga} = 180.000.000^2 = 3.24e + 16$$

$$\text{Kamar tidur} = 2^2 = 4$$

$$\text{Kamar mandi} = 1^2 = 1$$

$$\text{Garasi} = 1^2 = 1$$

$$\text{Luas tanah} = 36^2 = 1296$$

$$\text{Luas bangunan} = 50^2 = 2500$$

Apabila semua atribut sudah selesai dikuadratkan maka langkah selanjutnya adalah jumlahkan data lalu diakarkan untuk masing-masing atribut. Lakukan

HOME  
(HTTPS://LIBRARY.GUNADARMA.AC.ID/DEPC  
SYSTEM/)

PEMBAYARAN  
(HTTPS://LIBRARY.GUNADARMA.AC.ID/DEPC  
SYSTEM/PEMBAYARAN)

BIODATA DIRI  
(HTTPS://LIBRARY.GUNADARMA.AC.ID/DEPC  
SYSTEM/ANGGOTA)

PEMINJAMAN BUKU  
(HTTPS://LIBRARY.GUNADARMA.AC.ID/DEPC  
SYSTEM/ANGGOTA/PEMINJAMAN)

PENULISAN ILMIAH  
(HTTPS://LIBRARY.GUNADARMA.AC.ID/DEPC  
SYSTEM/ANGGOTA/PENULISAN)

FOTOCOPY  
(HTTPS://LIBRARY.GUNADARMA.AC.ID/DEPC  
SYSTEM/ANGGOTA/FOTOCOPY)

SUMBANG BUKU  
(HTTPS://LIBRARY.GUNADARMA.AC.ID/DEPC  
SYSTEM/ANGGOTA/SUMBANGBUKU)

KASUS & NOTIFIKASI  
(HTTPS://LIBRARY.GUNADARMA.AC.ID/DEPC  
SYSTEM/ANGGOTA/NOTIFIKASI)

BEBAS PERPUSTAKAAN  
(HTTPS://LIBRARY.GUNADARMA.AC.ID/DEPC  
SYSTEM/ANGGOTA/BEBASTERPUSTAKAAN)

KUNJUNGAN  
(HTTPS://LIBRARY.GUNADARMA.AC.ID/DEPC  
SYSTEM/KUNJUNGAN)

E-PAPER  
(HTTPS://LIBRARY.GUNADARMA.AC.ID/DEPC

perhitungan pada semua data sampel sehingga mendapatkan hasil untuk masing-masing data.

Tabel 3. 27 Hasil Perhitungan Kuadrat Bobot

	data(1)	data(2)	data(3)	data(4)	data(5)	data(6)	data(7)
daerah	0	0	64	0	64	0	0
harga	3.24E+16	3.99424E+17	7.0225E+16	3.56409E+17	9.025E+17	5.184E+17	3.6E+17
kamar tidur	4	4	4	4	9	4	4
kamar mandi	1	1	1	1	9	1	1
garasi	1	0	1	1	1	4	1
luas tanah	1296	4900	3025	2500	22500	4900	3600
luas bangunan	2500	2500	1296	1600	8100	3600	3136
SUM	3802	7405	4391	4106	30683	8509	6742
AKAR	61.66	86.05	66.26	64.08	175.17	92.24	82.11

Berdasarkan Tabel 3.27 di atas merupakan hasil perhitungan dari seluruh data sampel yang sudah diakarkan. Langkah selanjutnya mencari nilai *Cosine Similarity* yaitu membagi hasil penjumlahan perkalian *query* dan *item* dengan hasil akar kuadrat bobot sesuai dengan persamaan 2.3.

Data (1):

$$\text{Cosine} = \frac{4966}{61.66} = 80.5379$$

Tabel 3. 28 Hasil Nilai *Cosine Similarity*

	data(1)	data(2)	data(3)	data(4)	data(5)	data(6)
COSINE	80.53796	81.40397	80.31435	81.86887	80.20979	82.03222

Berdasarkan Tabel 3.28 di atas merupakan hasil dari perhitungan rekomendasi menggunakan *Cosine Similarity*. Pada hasil ini akan diberikan 3 rekomendasi dengan nilai cosine tertinggi yaitu data 6, data 4, dan data 2 merupakan urutan hasil rekomendasi yang akan diberikan kepada pengguna. Agar mendapatkan hasil rekomendasi yang digunakan dengan metode *content based filtering* dimana terdapat beberapa *filter* tambahan untuk menghasilkan rekomendasi yang lebih baik seperti penekanan pada fitur daerah, luas tanah dan

HOME  
(HTTPS://LIBRARY.GUNADARMA.AC.ID/DEPC  
SYSTEM/)

PEMBAYARAN  
(HTTPS://LIBRARY.GUNADARMA.AC.ID/DEPC  
SYSTEM/PEMBAYARAN)

BIODATA DIRI  
(HTTPS://LIBRARY.GUNADARMA.AC.ID/DEPC  
SYSTEM/ANGGOTA)

PEMINJAMAN BUKU  
(HTTPS://LIBRARY.GUNADARMA.AC.ID/DEPC  
SYSTEM/ANGGOTA/PEMINJAMAN)

PENULISAN ILMIAH  
(HTTPS://LIBRARY.GUNADARMA.AC.ID/DEPC  
SYSTEM/ANGGOTA/PENULISAN)

FOTOCOPY  
(HTTPS://LIBRARY.GUNADARMA.AC.ID/DEPC  
SYSTEM/ANGGOTA/FOTOCOPY)

SUMBANG BUKU  
(HTTPS://LIBRARY.GUNADARMA.AC.ID/DEPC  
SYSTEM/ANGGOTA/SUMBANGBUKU)

KASUS & NOTIFIKASI  
(HTTPS://LIBRARY.GUNADARMA.AC.ID/DEPC  
SYSTEM/ANGGOTA/NOTIFIKASI)

BEBAS PERPUSTAKAAN  
(HTTPS://LIBRARY.GUNADARMA.AC.ID/DEPC  
SYSTEM/ANGGOTA/BEBASTERPUSTAKAAN)

KUNJUNGAN  
(HTTPS://LIBRARY.GUNADARMA.AC.ID/DEPC  
SYSTEM/KUNJUNGAN)

E-PAPER  
(HTTPS://LIBRARY.GUNADARMA.AC.ID/DEPC

harga. Fitur tersebut akan menyesuaikan dengan kriteria rumah yang dipilih pengguna saat melakukan prediksi harga.

Tabel 3. 29 Data Hasil Rekomendasi

	pengguna	Rekomendasi		
		Rumah Bandung Mewah	Rumah Mezanin di Bandung	Rumah di Bandung
daerah	0	0	0	0
harga	600000000	632000000	597000000	720000000
kamar tidur	2	2	2	2
kamar mandi	2	1	1	1
garasi	1	0	1	2
luas tanah	60	70	50	70
luas bangunan	56	50	40	60

Berdasarkan Tabel 3.29 merupakan hasil rekomendasi yang diberikan berdasarkan data sampel yang diberikan. Pada hasil tersebut dapat dilihat bahwa dari data rekomendasi memiliki beberapa kemiripan seperti kota yang sama, jumlah kamar mandi yang sama dan sebagainya.

### 3.5. Evaluasi

Pada tahap ini akan dilakukan pengukuran untuk model yang telah dibentuk baik model prediksi maupun model sistem rekomendasi. Model yang sudah terbentuk akan diuji menggunakan 3 metrik untuk menentukan tipe kriteria pemisahan data terbaik pada model. Model yang mendapatkan 2 nilai metrik tertinggi akan dijadikan sebagai model terbaik kemudian akan dilanjutkan untuk penyimpanan pada tahap *deployment*. Metrik evaluasi yang digunakan yaitu *Mean Squared Error*, *Mean Absolute Error*, dan *R-Squared*.

- [HOME  
\(HTTPS://LIBRARY.GUNADARMA.AC.ID/DEPC  
SYSTEM/\)](#)
- [PEMBAYARAN  
\(HTTPS://LIBRARY.GUNADARMA.AC.ID/DEPC  
SYSTEM/PEMBAYARAN\)](#)
- [BIODATA DIRI  
\(HTTPS://LIBRARY.GUNADARMA.AC.ID/DEPC  
SYSTEM/ANGGOTA\)](#)
- [PEMINJAMAN BUKU  
\(HTTPS://LIBRARY.GUNADARMA.AC.ID/DEPC  
SYSTEM/ANGGOTA/PEMINJAMAN\)](#)
- [PENULISAN ILMIAH  
\(HTTPS://LIBRARY.GUNADARMA.AC.ID/DEPC  
SYSTEM/ANGGOTA/PENULISAN\)](#)
- [FOTOCOPY  
\(HTTPS://LIBRARY.GUNADARMA.AC.ID/DEPC  
SYSTEM/ANGGOTA/FOTOCOPY\)](#)
- [SUMBANG BUKU  
\(HTTPS://LIBRARY.GUNADARMA.AC.ID/DEPC  
SYSTEM/ANGGOTA/SUMBANGBUKU\)](#)
- [KASUS & NOTIFIKASI  
\(HTTPS://LIBRARY.GUNADARMA.AC.ID/DEPC  
SYSTEM/ANGGOTA/NOTIFIKASI\)](#)
- [BEBAS PERPUSTAKAAN  
\(HTTPS://LIBRARY.GUNADARMA.AC.ID/DEPC  
SYSTEM/ANGGOTA/BEBASPERPUSTAKAAN\)](#)
- [KUNJUNGAN  
\(HTTPS://LIBRARY.GUNADARMA.AC.ID/DEPC  
SYSTEM/KUNJUNGAN\)](#)
- [E-PAPER  
\(HTTPS://LIBRARY.GUNADARMA.AC.ID/DEPC](#)

### 3.5.1. MSE (Mean Squared Error)

Nilai metrik yang dihasilkan berupa nilai *error* yang dikuadratkan sehingga dapat menilai *error* yang terjadi pada saat model memprediksi hasil yang akan diberikan. Proses perhitungan manual akan menggunakan persamaan 2.12.

Tabel 3. 30 Data Perhitungan MSE *Random Forest Regression*

daerah	kamar tidur	kamar mandi	garasi	luas tanah	luas bangunan	harga	prediksi
8	2	1	1	66	45	268000000	357251000
2	3	1	1	75	108	900000000	899894000
0	3	3	0	95	106	1200000000	1383176000
0	3	3	2	111	200	1950000000	2146445000
8	3	3	1	120	110	1470000000	1563554000
0	3	2	1	93	104	1100000000	1169243500
0	2	1	0	60	40	200000000	217751500
8	2	1	2	75	45	585000000	626748500
3	2	2	1	60	30	330000000	469014500
2	4	2	1	170	100	1500000000	1616040000
8	3	2	0	89	103	1010000000	1109867500
8	3	1	0	89	90	880000000	918181500
8	1	2	2	63	61	230000000	255976000
2	1	2	0	55	30	168000000	235089500
0	4	3	0	190	180	2100000000	2370935000

Berdasarkan Tabel 3.30 merupakan hasil prediksi menggunakan data aktual.

Hasil prediksi harga yang diberikan tidak berbeda jauh dengan harga aktual maka dari itu harus dilakukan perhitungan *error* yang terjadi antara hasil prediksi dan data aktual. Langkah selanjutnya adalah melakukan proses penyamarataan rentang nilai menjadi 0-1 dengan menggunakan rumus *MinMaxScaler* pada persamaan 2.10.

$$\text{newdata} = \frac{(data - \min) \times (newmax - newmin)}{(\max - \min)} + newmin$$

$$\begin{aligned} \text{Harga}(1) &= \frac{(268000000 - 168000000) \times (1 - 0)}{(2100000000 - 168000000)} + 0 \\ &= \frac{100000000 \times 1}{1932000000} + 0 = 0,05176 \end{aligned}$$

Lakukan perhitungan minmax pada seluruh data kolom harga dan kolom hasil prediksi agar mendapatkan rentang nilai 0-1.

HOME  
(HTTPS://LIBRARY.GUNADARMA.AC.ID/DEPC  
SYSTEM/)

PEMBAYARAN  
(HTTPS://LIBRARY.GUNADARMA.AC.ID/DEPC  
SYSTEM/PEMBAYARAN)

BIODATA DIRI  
(HTTPS://LIBRARY.GUNADARMA.AC.ID/DEPC  
SYSTEM/ANGGOTA)

PEMINJAMAN BUKU  
(HTTPS://LIBRARY.GUNADARMA.AC.ID/DEPC  
SYSTEM/ANGGOTA/PEMINJAMAN)

PENULISAN ILMIAH  
(HTTPS://LIBRARY.GUNADARMA.AC.ID/DEPC  
SYSTEM/ANGGOTA/PENULISAN)

FOTOCOPY  
(HTTPS://LIBRARY.GUNADARMA.AC.ID/DEPC  
SYSTEM/ANGGOTA/FOTOCOPY)

SUMBANG BUKU  
(HTTPS://LIBRARY.GUNADARMA.AC.ID/DEPC  
SYSTEM/ANGGOTA/SUMBANGBUKU)

KASUS & NOTIFIKASI  
(HTTPS://LIBRARY.GUNADARMA.AC.ID/DEPC  
SYSTEM/ANGGOTA/NOTIFIKASI)

BEBAS PERPUSTAKAAN  
(HTTPS://LIBRARY.GUNADARMA.AC.ID/DEPC  
SYSTEM/ANGGOTA/BEBASPERPUSTAKAAN)

KUNJUNGAN  
(HTTPS://LIBRARY.GUNADARMA.AC.ID/DEPC  
SYSTEM/KUNJUNGAN)

E-PAPER  
(HTTPS://LIBRARY.GUNADARMA.AC.ID/DEPC

Tabel 3. 31 Hasil Data Proses Perhitungan MSE *Random Forest Regression*

harga	prediksi	minmax harga(X)	minmax pred(F)	X-F	(X-F) <sup>2</sup>
268000000	357251000	0.05176	0.06479	-0.01303	0.00017
900000000	899894000	0.37888	0.31681	0.06208	0.00385
1200000000	1383176000	0.53416	0.54126	-0.00709	0.00005
1950000000	2146445000	0.92236	0.89574	0.02662	0.00071
1470000000	1563554000	0.67391	0.62503	0.04888	0.00239
1100000000	1169243500	0.48240	0.44190	0.04050	0.00164
2000000000	217751500	0.01656	0.00000	0.01656	0.00027
585000000	626748500	0.21584	0.18995	0.02589	0.00067
330000000	469014500	0.08385	0.11669	-0.03284	0.00108
1500000000	1616040000	0.68944	0.64941	0.04004	0.00160
1010000000	1109867500	0.43582	0.41432	0.02149	0.00046
880000000	918181500	0.36853	0.32530	0.04323	0.00187
230000000	255976000	0.03209	0.01775	0.01434	0.00021
168000000	235089500	0.00000	0.00805	-0.00805	0.00006
2100000000	2370935000	1.00000	1.00000	0.00000	0.00000

Berdasarkan Tabel 3.31 di atas pada kolom X-F merupakan hasil pengurangan antara nilai *minmax* kolom harga dengan nilai *minmax* kolom prediksi. Hasil dari pengurangan tersebut dikuadratkan dan diletakkan pada kolom (X-F)<sup>2</sup>. Jika sudah hitung nilai MSE menggunakan persamaan 2.12.

$$MSE = \frac{1}{n} \sum_{i=1}^n (X_i - F_i)^2$$

$$MSE = \frac{0.00017 + 0.00385 + 0.00005 + 0.00071 + \dots + 0.00006 + 0.00000}{15}$$

$$MSE = \frac{0.01504}{15} = 0.00102 = 0.102\%$$

Nilai MSE yang berhasil di dapatkan sebesar 0.00102 jika di jadikan ke dalam bentuk persen maka menghasilkan nilai 0.102%. Berdasarkan nilai yang di dapatkan bahwa nilai *error* untuk 15 data sampel yang terjadi pada model cukup kecil dan model yang berhasil dibentuk dengan dataset sudah baik dalam melakukan prediksi.

Langkah selanjutnya akan dilakukan perhitungan manual untuk evaluasi hasil sistem rekomendasi yang diberikan. Data yang akan digunakan merupakan 3 data hasil prediksi pada proses *modeling*. Dikarenakan target data adalah kolom

HOME  
(HTTPS://LIBRARY.GUNADARMA.AC.ID/DEPC  
SYSTEM/)

PEMBAYARAN  
(HTTPS://LIBRARY.GUNADARMA.AC.ID/DEPC  
SYSTEM/PEMBAYARAN)

BIODATA DIRI  
(HTTPS://LIBRARY.GUNADARMA.AC.ID/DEPC  
SYSTEM/ANGGOTA)

PEMINJAMAN BUKU  
(HTTPS://LIBRARY.GUNADARMA.AC.ID/DEPC  
SYSTEM/ANGGOTA/PEMINJAMAN)

PENULISAN ILMIAH  
(HTTPS://LIBRARY.GUNADARMA.AC.ID/DEPC  
SYSTEM/ANGGOTA/PENULISAN)

FOTOCOPY  
(HTTPS://LIBRARY.GUNADARMA.AC.ID/DEPC  
SYSTEM/ANGGOTA/FOTOCOPY)

SUMBANG BUKU  
(HTTPS://LIBRARY.GUNADARMA.AC.ID/DEPC  
SYSTEM/ANGGOTA/SUMBANGBUKU)

KASUS & NOTIFIKASI  
(HTTPS://LIBRARY.GUNADARMA.AC.ID/DEPC  
SYSTEM/ANGGOTA/NOTIFIKASI)

BEBAS PERPUSTAKAAN  
(HTTPS://LIBRARY.GUNADARMA.AC.ID/DEPC  
SYSTEM/ANGGOTA/BEBASTERPUSTAKAAN)

KUNJUNGAN  
(HTTPS://LIBRARY.GUNADARMA.AC.ID/DEPC  
SYSTEM/KUNJUNGAN)

E-PAPER  
(HTTPS://LIBRARY.GUNADARMA.AC.ID/DEPC

harga maka perhitungan akan dilakukan dengan membandingkan hasil harga prediksi rekomendasi dengan harga rekomendasi yang sesuai dengan profil pengguna.

Tabel 3. 32 Data Perhitungan MSE *Cosine Similarity*

	daerah	kamar tidur	kamar mandi	garasi	luas tanah	luas bangunan	harga	prediksi
Rumah di Bandung	0	2	1	0	70	50	632000000	642000000
Rumah Mezanin di Bandung	0	2	1	1	50	40	597000000	562000000
Rumah Bandung Mewah	0	2	1	2	70	60	720000000	707000000

Berdasarkan Tabel 3.32 di atas pada kolom prediksi merupakan hasil yang didapatkan dari prediksi menggunakan *Cosine Similarity* dengan data pengujian. Kemudian data ini akan diubah menjadi rentang 0-1 menggunakan perhitungan *minmax* seperti sebelumnya.

Tabel 3. 33 Hasil Data Proses Perhitungan MSE *Cosine Similarity*

harga	prediksi	minmax harga (x)	minmax pred (f)	X-F	(X-F)^2
632000000.00	642000000.00	0.28455	0.55172	-0.26717	0.071380499
597000000.00	562000000.00	0.00000	0.00000	0.00000	0
720000000.00	707000000.00	1.00000	1.00000	0.00000	0

Berdasarkan Tabel 3.33 di atas adalah hasil pengubahan rentang nilai pada data menjadi 0-1. Langkah selanjutnya akan dilakukan perhitungan pada masing-masing data menggunakan rumus MSE dengan Persamaan 2.12.

$$MSE = \frac{1}{n} \sum_{i=1}^n (X_i - F_i)^2$$

$$MSE = \frac{0.071380499+0+0}{3}$$

HOME  
(HTTPS://LIBRARY.GUNADARMA.AC.ID/DEPC  
SYSTEM/)

PEMBAYARAN  
(HTTPS://LIBRARY.GUNADARMA.AC.ID/DEPC  
SYSTEM/PEMBAYARAN)

BIODATA DIRI  
(HTTPS://LIBRARY.GUNADARMA.AC.ID/DEPC  
SYSTEM/ANGGOTA)

PEMINJAMAN BUKU  
(HTTPS://LIBRARY.GUNADARMA.AC.ID/DEPC  
SYSTEM/ANGGOTA/PEMINJAMAN)

PENULISAN ILMIAH  
(HTTPS://LIBRARY.GUNADARMA.AC.ID/DEPC  
SYSTEM/ANGGOTA/PENULISAN)

FOTOCOPY  
(HTTPS://LIBRARY.GUNADARMA.AC.ID/DEPC  
SYSTEM/ANGGOTA/FOTOCOPY)

SUMBANG BUKU  
(HTTPS://LIBRARY.GUNADARMA.AC.ID/DEPC  
SYSTEM/ANGGOTA/SUMBANGBUKU)

KASUS & NOTIFIKASI  
(HTTPS://LIBRARY.GUNADARMA.AC.ID/DEPC  
SYSTEM/ANGGOTA/NOTIFIKASI)

BEBAS PERPUSTAKAAN  
(HTTPS://LIBRARY.GUNADARMA.AC.ID/DEPC  
SYSTEM/ANGGOTA/BEBASPERPUSTAKAAN)

KUNJUNGAN  
(HTTPS://LIBRARY.GUNADARMA.AC.ID/DEPC  
SYSTEM/KUNJUNGAN)

E-PAPER  
(HTTPS://LIBRARY.GUNADARMA.AC.ID/DEPC

$$MSE = \frac{0.071380499}{3} = 0.0237935$$

Hasil MSE pada sistem rekomendasi menggunakan algoritma *Cosine Similarity* adalah 0.0237 dan dapat dikatakan bahwa algoritma sudah memberikan hasil rekomendasi yang baik pada sistem.

### 3.5.2. MAE (*Mean Absolute Error*)

Nilai metrik ini akan menghasilkan nilai *error absolute* dari hasil prediksi yang diberikan oleh model. Proses perhitungan manual akan menggunakan Persamaan 2.13. Pada perhitungan *Mean Absolute Error* akan menggunakan 15 data sampel.

Tabel 3. 34 Hasil Data Proses Perhitungan MAE *Random Forest Regression*

harga	prediksi	minmaxharga(X)	minmaxpred(F)	X-F	X-F
268000000	357251000	0.05176	0.06479	-0.01303	0.01303
900000000	899894000	0.37888	0.31681	0.06208	0.06208
1200000000	1383176000	0.53416	0.54126	-0.00709	0.00709
1950000000	2146445000	0.92236	0.89574	0.02662	0.02662
1470000000	1563554000	0.67391	0.62503	0.04888	0.04888
1100000000	1169243500	0.48240	0.44190	0.04050	0.04050
2000000000	217751500	0.01656	0.00000	0.01656	0.01656
585000000	626748500	0.21584	0.18995	0.02589	0.02589
330000000	469014500	0.08385	0.11669	-0.03284	0.03284
1500000000	1616040000	0.68944	0.64941	0.04004	0.04004
1010000000	1109867500	0.43582	0.41432	0.02149	0.02149
880000000	918181500	0.36853	0.32530	0.04323	0.04323
230000000	255976000	0.03209	0.01775	0.01434	0.01434
168000000	235089500	0.00000	0.00805	-0.00805	0.00805
2100000000	2370935000	1.00000	1.00000	0.00000	0.00000

Berdasarkan Tabel 3.33 di atas merupakan hasil proses dari perhitungan MAE pada algoritma *Random Forest Regression*. Masukan nilai pada kolom absolute X-F ke dalam rumus MAE pada persamaan 2.13.

$$MAE = \frac{1}{n} \sum_{j=1}^n |y_j - \hat{y}_j|$$

HOME  
(HTTPS://LIBRARY.GUNADARMA.AC.ID/DEPC  
SYSTEM/)

PEMBAYARAN  
(HTTPS://LIBRARY.GUNADARMA.AC.ID/DEPC  
SYSTEM/PEMBAYARAN)

BIODATA DIRI  
(HTTPS://LIBRARY.GUNADARMA.AC.ID/DEPC  
SYSTEM/ANGGOTA)

PEMINJAMAN BUKU  
(HTTPS://LIBRARY.GUNADARMA.AC.ID/DEPC  
SYSTEM/ANGGOTA/PEMINJAMAN)

PENULISAN ILMIAH  
(HTTPS://LIBRARY.GUNADARMA.AC.ID/DEPC  
SYSTEM/ANGGOTA/PENULISAN)

FOTOCOPY  
(HTTPS://LIBRARY.GUNADARMA.AC.ID/DEPC  
SYSTEM/ANGGOTA/FOTOCOPY)

SUMBANG BUKU  
(HTTPS://LIBRARY.GUNADARMA.AC.ID/DEPC  
SYSTEM/ANGGOTA/SUMBANGBUKU)

KASUS & NOTIFIKASI  
(HTTPS://LIBRARY.GUNADARMA.AC.ID/DEPC  
SYSTEM/ANGGOTA/NOTIFIKASI)

BEBAS PERPUSTAKAAN  
(HTTPS://LIBRARY.GUNADARMA.AC.ID/DEPC  
SYSTEM/ANGGOTA/BEBASPERPUSTAKAAN)

KUNJUNGAN  
(HTTPS://LIBRARY.GUNADARMA.AC.ID/DEPC  
SYSTEM/KUNJUNGAN)

E-PAPER  
(HTTPS://LIBRARY.GUNADARMA.AC.ID/DEPC

$$MAE = \frac{0.01303+0.06208+0.00709+\dots+0.00805+0.00000}{15}$$

$$MAE = \frac{0.40065}{15} = 0.02671 = 2.671\%$$

Hasil perhitungan manual MSE pada algoritma *Random Forest* dengan 15 data sampel sebesar 0.02925 jika dijadikan persen sebesar 2.925% sehingga dapat dikatakan bahwa model dapat memberikan prediksi harga cukup baik. Selanjutnya akan dilakukan perhitungan manual untuk hasil rekomendasi pada sistem.

Tabel 3. 35 Hasil Proses Perhitungan MAE *Cosine Similarity*

harga	prediksi	minmax aktual (x)	minmax pred (f)	X-F	X-F
632000000.00	642000000.00	0.28455	0.55172	-0.26717	0.267171292
597000000.00	562000000.00	0.00000	0.00000	0.00000	0
720000000.00	707000000.00	1.00000	1.00000	0.00000	0

Berdasarkan Tabel 3.35 pada kolom |X-F| adalah hasil perhitungan nilai absolute dari pengurangan data aktual dengan data prediksi. Langkah selanjutnya adalah perhitungan MAE dengan Persamaan 2.13.

$$MAE = \frac{1}{n} \sum_{j=1}^n |y_j - \hat{y}_j|$$

$$MAE = \frac{0.267171292+0+0}{3}$$

$$MAE = \frac{0.267171292}{3} = 0.089057097$$

Hasil perhitungan manual MAE pada algoritma *Cosine Similarity* dengan data sampel sebesar 0.089 atau 89% sehingga dapat dikatakan bahwa rekomendasi yang diberikan oleh sistem sudah baik.

### 3.5.3. *R-Squared*

Nilai metrik yang dihasilkan cukup berbeda dengan metrik evaluasi sebelumnya. Pada pengukuran evaluasi ini akan melihat seberapa besar variabel target di pengaruhi oleh variabel independennya. Jenis evaluasi ini juga dapat dijadikan sebagai akurasi dari keterikatan variabel yang telah memberikan hasil prediksi. Proses perhitungan manual akan menggunakan Persamaan 2.14

HOME  
(HTTPS://LIBRARY.GUNADARMA.AC.ID/DEPC  
SYSTEM/)

PEMBAYARAN  
(HTTPS://LIBRARY.GUNADARMA.AC.ID/DEPC  
SYSTEM/PEMBAYARAN)

BIODATA DIRI  
(HTTPS://LIBRARY.GUNADARMA.AC.ID/DEPC  
SYSTEM/ANGGOTA)

PEMINJAMAN BUKU  
(HTTPS://LIBRARY.GUNADARMA.AC.ID/DEPC  
SYSTEM/ANGGOTA/PEMINJAMAN)

PENULISAN ILMIAH  
(HTTPS://LIBRARY.GUNADARMA.AC.ID/DEPC  
SYSTEM/ANGGOTA/PENULISAN)

FOTOCOPY  
(HTTPS://LIBRARY.GUNADARMA.AC.ID/DEPC  
SYSTEM/ANGGOTA/FOTOCOPY)

SUMBANG BUKU  
(HTTPS://LIBRARY.GUNADARMA.AC.ID/DEPC  
SYSTEM/ANGGOTA/SUMBANGBUKU)

KASUS & NOTIFIKASI  
(HTTPS://LIBRARY.GUNADARMA.AC.ID/DEPC  
SYSTEM/ANGGOTA/NOTIFIKASI)

BEBAS PERPUSTAKAAN  
(HTTPS://LIBRARY.GUNADARMA.AC.ID/DEPC  
SYSTEM/ANGGOTA/BEBASTERPUSTAKAAN)

KUNJUNGAN  
(HTTPS://LIBRARY.GUNADARMA.AC.ID/DEPC  
SYSTEM/KUNJUNGAN)

E-PAPER  
(HTTPS://LIBRARY.GUNADARMA.AC.ID/DEPC

menggunakan 15 data sampel. Selanjutnya akan dilakukan perhitungan SSE dan SST menggunakan rumus pada Persamaan 2.14.

$$SSE = \sum(Y_{aktual} - Y_{pred})^2$$

$$SSE = ((268000000 - 357251000)^2 + \dots + (2100000000 - 2370935000)^2)$$

$$SSE = (7.9657E + 15) + (11236000000) + \dots + (7.34058E + 16)$$

$$SSE = 2.18518E + 17$$

$$\bar{Y} = \frac{268000000+900000000+\dots+2100000000}{15} = 926066666.7$$

$$SST = \sum(Y_{aktual} - \bar{Y})^2$$

$$SST = ((268000000 - 926066666.7)^2 + \dots + (2100000000 - 926066666.7)^2)$$

$$SST = (4.3305E + 17) + (6.79471E + 14) + \dots + (1.37812E + 18)$$

$$SST = 5.65798E + 18$$

Tabel 3. 36 Hasil Perhitungan SSE dan SST *Random Forest Regression*

	Harga	Prediksi	SSE	SST
	268000000	357251000	7.96574E+15	3.55295E+17
	900000000	899894000	11236000000	6.79471E+14
	1200000000	1383176000	3.35534E+16	7.50395E+16
	1950000000	2146445000	3.85906E+16	1.04844E+18
	1470000000	1563554000	8.75235E+15	2.95863E+17
	1100000000	1169243500	4.79466E+15	3.02528E+16
	2000000000	217751500	3.15116E+14	5.27173E+17
	585000000	626748500	1.74294E+15	1.16326E+17
	330000000	469014500	1.9325E+16	4.33052E+17
	1500000000	1616040000	1.34653E+16	3.29399E+17
	1010000000	1109867500	9.97352E+15	7.0448E+15
	880000000	918181500	1.45783E+15	2.12214E+15
	230000000	255976000	6.74753E+14	4.84509E+17
	168000000	235089500	4.501E+15	5.74665E+17
	2100000000	2370935000	7.34058E+16	1.37812E+18
Rata-rata	926066666.7	Total	2.18518E+17	5.65798E+18

HOME  
(HTTPS://LIBRARY.GUNADARMA.AC.ID/DEPC  
SYSTEM/)

PEMBAYARAN  
(HTTPS://LIBRARY.GUNADARMA.AC.ID/DEPC  
SYSTEM/PEMBAYARAN)

BIODATA DIRI  
(HTTPS://LIBRARY.GUNADARMA.AC.ID/DEPC  
SYSTEM/ANGGOTA)

PEMINJAMAN BUKU  
(HTTPS://LIBRARY.GUNADARMA.AC.ID/DEPC  
SYSTEM/ANGGOTA/PEMINJAMAN)

PENULISAN ILMIAH  
(HTTPS://LIBRARY.GUNADARMA.AC.ID/DEPC  
SYSTEM/ANGGOTA/PENULISAN)

FOTOCOPY  
(HTTPS://LIBRARY.GUNADARMA.AC.ID/DEPC  
SYSTEM/ANGGOTA/FOTOCOPY)

SUMBANG BUKU  
(HTTPS://LIBRARY.GUNADARMA.AC.ID/DEPC  
SYSTEM/ANGGOTA/SUMBANGBUKU)

KASUS & NOTIFIKASI  
(HTTPS://LIBRARY.GUNADARMA.AC.ID/DEPC  
SYSTEM/ANGGOTA/NOTIFIKASI)

BEBAS PERPUSTAKAAN  
(HTTPS://LIBRARY.GUNADARMA.AC.ID/DEPC  
SYSTEM/ANGGOTA/BEBASPERPUSTAKAAN)

KUNJUNGAN  
(HTTPS://LIBRARY.GUNADARMA.AC.ID/DEPC  
SYSTEM/KUNJUNGAN)

E-PAPER  
(HTTPS://LIBRARY.GUNADARMA.AC.ID/DEPC

Berdasarkan Tabel 3.36 di atas merupakan hasil perhitungan SSE dan SSR menggunakan 15 data sampel. Hasil dari perhitungan tersebut akan dimasukan ke dalam perhitungan *R-Squared* pada persamaan 2.14.

$$R^2 = \frac{SSR}{SST} = 1 - \frac{SSE}{SST} = 1 - \frac{\sum(Y_{aktual} - Y_{pred})^2}{\sum(Y_{aktual} - \bar{Y})^2}$$

$$R^2 = 1 - \frac{2.18518E+17}{5.65798E+18} = 0.96137878$$

Hasil perhitungan *R-Squared* yang diperoleh secara manual dengan 15 data sampel sebesar 0.9614 sedangkan untuk perhitungan rekomendasi dengan algoritma *Cosine Similarity* sebagai berikut.

$$SSE = \sum(Y_{aktual} - Y_{pred})^2$$

$$SSE = ((632000000 - 642000000)^2 + (597000000 - 562000000)^2 + (720000000 - 707000000)^2)$$

$$SSE = (1E + 14) + (11.225E + 15) + (1.69E + 14)$$

$$SSE = 1.494E + 15$$

$$\bar{Y} = \frac{632000000 + 597000000 + 720000000}{3} = 649666666.67$$

$$SST = \sum(Y_{aktual} - \bar{Y})^2$$

$$SST = ((632000000 - 649666666.67)^2 + (597000000 - 649666666.67)^2 + (720000000 - 649666666.67)^2)$$

$$SST = (3.12111E + 14) + (2.77378E + 15) + (4.94678E+15)$$

$$SST = 8.03267E + 15$$

Tabel 3. 37 Hasil Perhitungan SSE dan SST *Cosine Similarity*

	Harga	Prediksi	SSE	SST
	632000000.00	642000000.00	1E+14	3.12111E+14
	597000000.00	562000000.00	1.225E+15	2.77378E+15
	720000000.00	707000000.00	1.69E+14	4.94678E+15
Rata-Rata	649666666.67	Total	1.494E+15	8.03267E+15

HOME  
(HTTPS://LIBRARY.GUNADARMA.AC.ID/DEPC  
SYSTEM/)

PEMBAYARAN  
(HTTPS://LIBRARY.GUNADARMA.AC.ID/DEPC  
SYSTEM/PEMBAYARAN)

BIODATA DIRI  
(HTTPS://LIBRARY.GUNADARMA.AC.ID/DEPC  
SYSTEM/ANGGOTA)

PEMINJAMAN BUKU  
(HTTPS://LIBRARY.GUNADARMA.AC.ID/DEPC  
SYSTEM/ANGGOTA/PEMINJAMAN)

PENULISAN ILMIAH  
(HTTPS://LIBRARY.GUNADARMA.AC.ID/DEPC  
SYSTEM/ANGGOTA/PENULISAN)

FOTOCOPY  
(HTTPS://LIBRARY.GUNADARMA.AC.ID/DEPC  
SYSTEM/ANGGOTA/FOTOCOPY)

SUMBANG BUKU  
(HTTPS://LIBRARY.GUNADARMA.AC.ID/DEPC  
SYSTEM/ANGGOTA/SUMBANGBUKU)

KASUS & NOTIFIKASI  
(HTTPS://LIBRARY.GUNADARMA.AC.ID/DEPC  
SYSTEM/ANGGOTA/NOTIFIKASI)

BEBAS PERPUSTAKAAN  
(HTTPS://LIBRARY.GUNADARMA.AC.ID/DEPC  
SYSTEM/ANGGOTA/BEBAZPERPUSTAKAAN)

KUNJUNGAN  
(HTTPS://LIBRARY.GUNADARMA.AC.ID/DEPC  
SYSTEM/KUNJUNGAN)

E-PAPER  
(HTTPS://LIBRARY.GUNADARMA.AC.ID/DEPC

Berdasarkan Tabel 3.37 di atas merupakan hasil perhitungan SSE dan SST yang kemudian akan dimasukan ke dalam rumus *R-Squared* pada Persamaan 2.14.

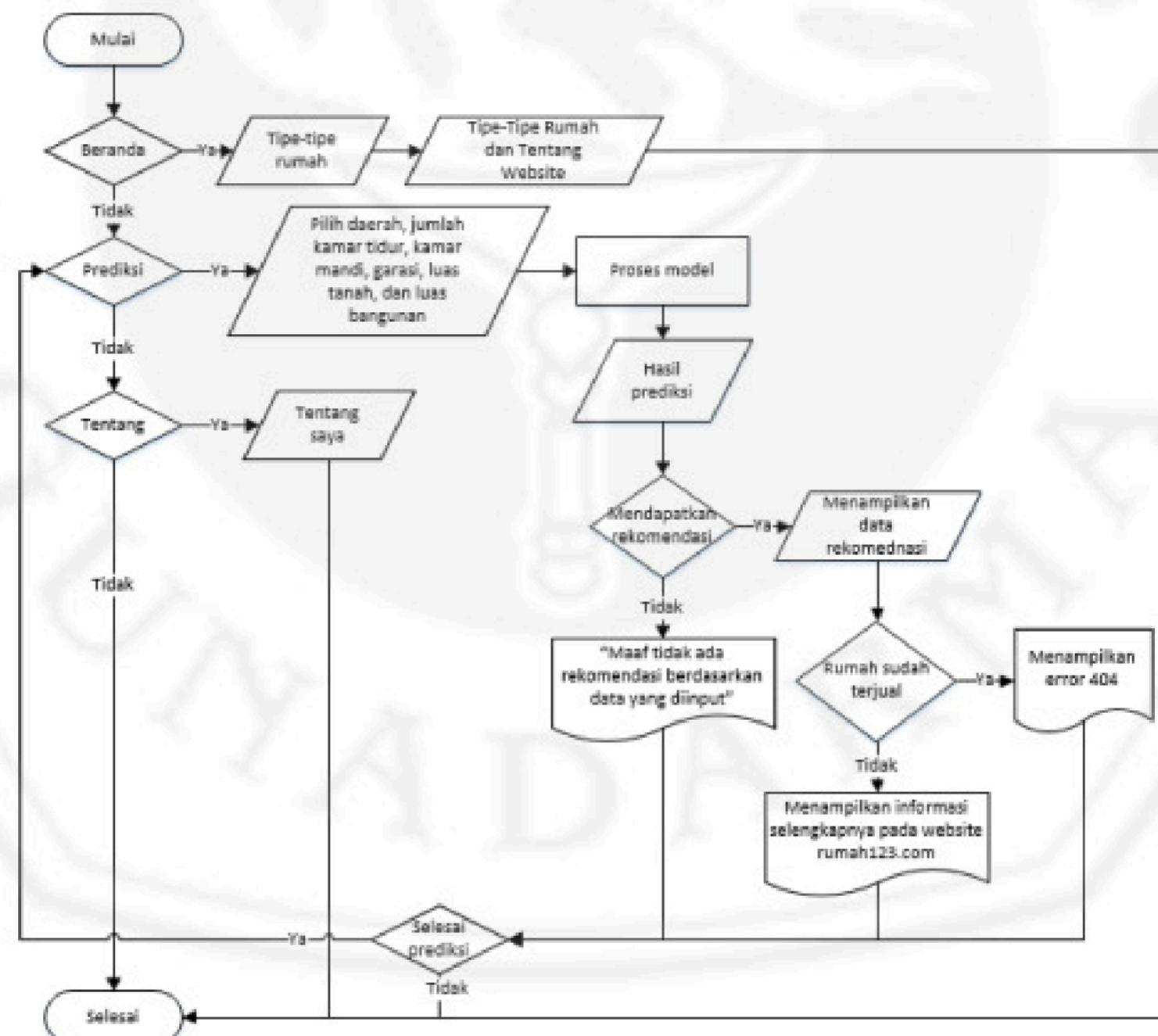
$$R^2 = \frac{SSR}{SST} = 1 - \frac{SSE}{SST} = 1 - \frac{\sum(Y_{aktual} - Y_{pred})^2}{\sum(Y_{aktual} - \bar{Y})^2}$$

$$R^2 = 1 - \frac{1.494E+15}{8.03267E+15} = 0.814009461$$

Hasil perhitungan *R-Squared* pada sistem rekomendasi menggunakan algoritma *Cosine Similarity* dengan 3 data sampel diperoleh sebesar 0.8140.

### 3.6. Deployment

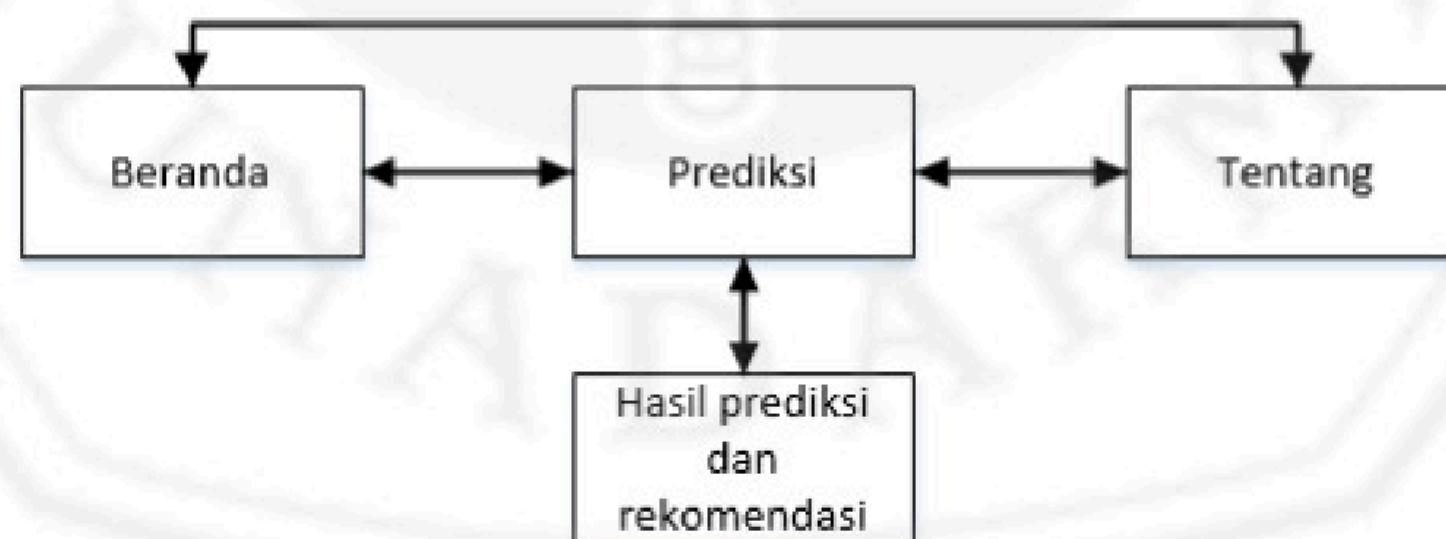
Pada proses terakhir adalah melakukan *deployment* terhadap model hasil *tuning parameter*. Pada penelitian ini akan melakukan *deployment* ke dalam sistem berbasis *website* dengan menggunakan *framework flask* pada bahasa pemrograman *Python*.



Gambar 3. 19 Flowchart Website Prediksi

- [HOME  
\(HTTPS://LIBRARY.GUNADARMA.AC.ID/DEPC/HOME/\)](https://library.gunadarma.ac.id/DEPC/HOME)
- [PEMBAYARAN  
\(HTTPS://LIBRARY.GUNADARMA.AC.ID/DEPC/SYSTEM/PEMBAYARAN\)](https://library.gunadarma.ac.id/DEPC/PEMBAYARAN)
- [BIODATA DIRI  
\(HTTPS://LIBRARY.GUNADARMA.AC.ID/DEPC/SYSTEM/ANGGOTA\)](https://library.gunadarma.ac.id/DEPC/BIODATA_DIRI)
- [PEMINJAMAN BUKU  
\(HTTPS://LIBRARY.GUNADARMA.AC.ID/DEPC/SYSTEM/ANGGOTA/PEMINJAMAN\)](https://library.gunadarma.ac.id/DEPC/PEMINJAMAN_BUKU)
- [PENULISAN ILMIAH  
\(HTTPS://LIBRARY.GUNADARMA.AC.ID/DEPC/SYSTEM/ANGGOTA/PENULISAN\)](https://library.gunadarma.ac.id/DEPC/PENULISAN_ILMIAH)
- [FOTOCOPY  
\(HTTPS://LIBRARY.GUNADARMA.AC.ID/DEPC/SYSTEM/ANGGOTA/FOTOCOPY\)](https://library.gunadarma.ac.id/DEPC/FOTOCOPY)
- [SUMBANG BUKU  
\(HTTPS://LIBRARY.GUNADARMA.AC.ID/DEPC/SYSTEM/ANGGOTA/SUMBANGBUKU\)](https://library.gunadarma.ac.id/DEPC/SUMBANG_BUKU)
- [KASUS & NOTIFIKASI  
\(HTTPS://LIBRARY.GUNADARMA.AC.ID/DEPC/SYSTEM/ANGGOTA/NOTIFIKASI\)](https://library.gunadarma.ac.id/DEPC/KASUS_NOTIFIKASI)
- [BEBAS PERPUSTAKAAN  
\(HTTPS://LIBRARY.GUNADARMA.AC.ID/DEPC/SYSTEM/ANGGOTA/BEBAСПЕРПУСТАКААН\)](https://library.gunadarma.ac.id/DEPC/BEBAS_PERPUSTAKAAN)
- [KUNJUNGAN  
\(HTTPS://LIBRARY.GUNADARMA.AC.ID/DEPC/SYSTEM/KUNJUNGAN\)](https://library.gunadarma.ac.id/DEPC/KUNJUNGAN)
- [E-PAPER  
\(HTTPS://LIBRARY.GUNADARMA.AC.ID/DEPC/EPAPER\)](https://library.gunadarma.ac.id/DEPC/E_PAPER)

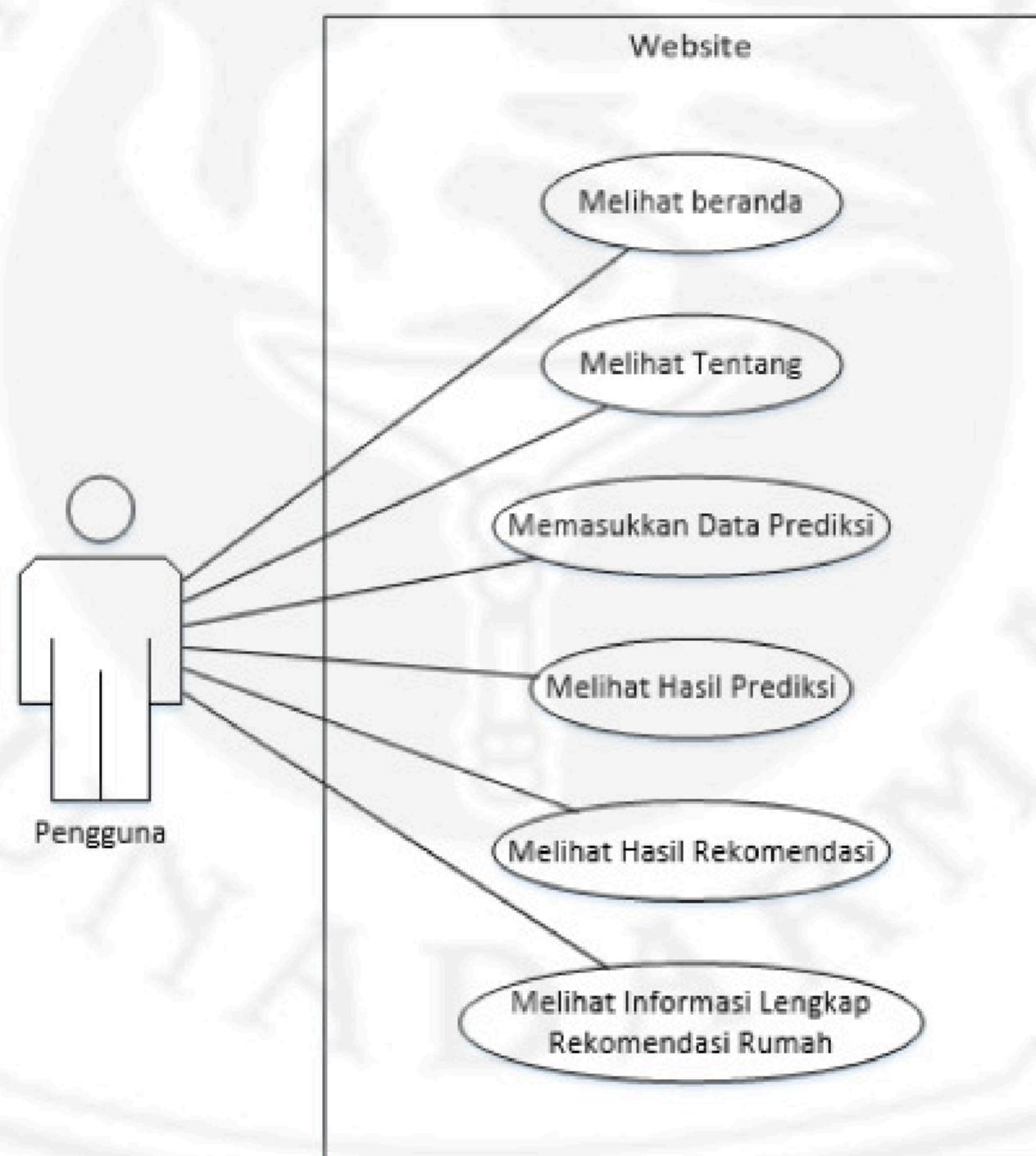
Berdasarkan Gambar 3.19 di atas dapat dilihat bahwa *website* ini pertama-tama berada pada menu beranda yang akan menampilkan informasi tipe-tipe rumah dan penjelasan singkat mengenai *website* prediksi tersebut. Jika tidak memilih halaman beranda dan memilih halaman prediksi maka pengguna akan memasukkan beberapa inputan yang dibutuhkan seperti daerah, kamar tidur, kamar mandi dan lainnya lalu inputan ini akan diproses. Pada proses ini model akan mempelajari data baru dengan pola yang baru kemudian akan menghasilkan prediksi, jika sudah diprediksi maka data akan dimasukan ke dalam sistem rekomendasi sehingga memberikan rekomendasi yang memiliki kemiripan dengan hasil prediksi. Proses selanjutnya akan menampilkan hasil prediksi kemudian apabila data yang dimasukkan mendapatkan rekomendasi maka akan menampilkan data rekomendasi dengan informasi selengkapnya pada sistem, akan tetapi apabila data yang dimasukkan tidak ada pada data rekomendasi yang sudah disimpan maka akan menampilkan “maaf tidak ada rekomendasi berdasarkan data yang di input”. Tombol informasi selengkapnya akan menampilkan informasi pada *website* rumah123.com jika rumah masih tersedia pada *website* tersebut, namun apabila rumah sudah terjual maka akan menampilkan *error* 404 pada *website* karena alamat URL sudah tidak di iklankan kembali. Jika pengguna menekan tombol selesai maka akan kembali ke halaman prediksi namun apabila tidak menekan tombol selesai maka akan keluar dari aplikasi dan proses selesai. Jika tidak memilih menu prediksi dan memilih menu tentang maka akan menampilkan informasi tentang penulis.



Gambar 3. 20 Struktur Navigasi *Website*

- [HOME  
\(HTTPS://LIBRARY.GUNADARMA.AC.ID/DEPC SYSTEM/\)](#)
- [PEMBAYARAN  
\(HTTPS://LIBRARY.GUNADARMA.AC.ID/DEPC SYSTEM/PEMBAYARAN\)](#)
- [BIODATA DIRI  
\(HTTPS://LIBRARY.GUNADARMA.AC.ID/DEPC SYSTEM/ANGGOTA\)](#)
- [PEMINJAMAN BUKU  
\(HTTPS://LIBRARY.GUNADARMA.AC.ID/DEPC SYSTEM/ANGGOTA/PEMINJAMAN\)](#)
- [PENULISAN ILMIAH  
\(HTTPS://LIBRARY.GUNADARMA.AC.ID/DEPC SYSTEM/ANGGOTA/PENULISAN\)](#)
- [FOTOCOPY  
\(HTTPS://LIBRARY.GUNADARMA.AC.ID/DEPC SYSTEM/ANGGOTA/FOTOCOPY\)](#)
- [SUMBANG BUKU  
\(HTTPS://LIBRARY.GUNADARMA.AC.ID/DEPC SYSTEM/ANGGOTA/SUMBANGBUKU\)](#)
- [KASUS & NOTIFIKASI  
\(HTTPS://LIBRARY.GUNADARMA.AC.ID/DEPC SYSTEM/ANGGOTA/NOTIFIKASI\)](#)
- [BEBAS PERPUSTAKAAN  
\(HTTPS://LIBRARY.GUNADARMA.AC.ID/DEPC SYSTEM/ANGGOTA/BEBASPERPUSTAKAAN\)](#)
- [KUNJUNGAN  
\(HTTPS://LIBRARY.GUNADARMA.AC.ID/DEPC SYSTEM/KUNJUNGAN\)](#)
- [E-PAPER  
\(HTTPS://LIBRARY.GUNADARMA.AC.ID/DEPC](#)

Berdasarkan Gambar 3.20 di atas merupakan jenis struktur navigasi campuran (*composite*) karena *website* yang akan dibangun memiliki struktur *linear* dan *non-linear* dimana pada halaman beranda dapat langsung menuju halaman tentang, kemudian dari halaman tentang dapat langsung memilih halaman prediksi, akan tetapi ada 1 halaman yang bersifat *non-linear* yaitu halaman hasil hanya dapat di akses jika melalui halaman prediksi. Pada halaman beranda akan menampilkan informasi tipe rumah dan penjelasan singkat tentang *website*. Kemudian pada halaman tentang akan menampilkan informasi mengenai tentang penulis. Pada halaman prediksi jika sudah memasukkan data yang diminta maka akan menampilkan hasil prediksi dan rekomendasi berdasarkan prediksi yang dihasilkan.



Gambar 3. 21 Use case Website

HOME  
(HTTPS://LIBRARY.GUNADARMA.AC.ID/DEPC  
SYSTEM/)

PEMBAYARAN  
(HTTPS://LIBRARY.GUNADARMA.AC.ID/DEPC  
SYSTEM/PEMBAYARAN)

BIODATA DIRI  
(HTTPS://LIBRARY.GUNADARMA.AC.ID/DEPC  
SYSTEM/ANGGOTA)

PEMINJAMAN BUKU  
(HTTPS://LIBRARY.GUNADARMA.AC.ID/DEPC  
SYSTEM/ANGGOTA/PEMINJAMAN)

PENULISAN ILMIAH  
(HTTPS://LIBRARY.GUNADARMA.AC.ID/DEPC  
SYSTEM/ANGGOTA/PENULISAN)

FOTOCOPY  
(HTTPS://LIBRARY.GUNADARMA.AC.ID/DEPC  
SYSTEM/ANGGOTA/FOTOCOPY)

SUMBANG BUKU  
(HTTPS://LIBRARY.GUNADARMA.AC.ID/DEPC  
SYSTEM/ANGGOTA/SUMBANGBUKU)

KASUS & NOTIFIKASI  
(HTTPS://LIBRARY.GUNADARMA.AC.ID/DEPC  
SYSTEM/ANGGOTA/NOTIFIKASI)

BEBAS PERPUSTAKAAN  
(HTTPS://LIBRARY.GUNADARMA.AC.ID/DEPC  
SYSTEM/ANGGOTA/BEBASPERPUSTAKAAN)

KUNJUNGAN  
(HTTPS://LIBRARY.GUNADARMA.AC.ID/DEPC  
SYSTEM/KUNJUNGAN)

E-PAPER  
(HTTPS://LIBRARY.GUNADARMA.AC.ID/DEPC

Berdasarkan Gambar 3.21 di atas merupakan *use case* dari *website* yang akan dibangun, *user* yang dapat menggunakan *website* ini hanya pengguna. Kegiatan yang dapat dilakukan oleh pengguna *website* meliputi melihat beranda yang menampilkan informasi *website* dan tipe-tipe rumah, kemudian dapat melihat halaman tentang yang menampilkan deskripsi singkat pembuat *website*. Pengguna juga dapat memasukkan data prediksi yang akan menghasilkan nilai prediksi dan rekomendasi dari implementasi model. Kegiatan lain yang dapat dilakukan pengguna adalah melihat hasil prediksi dan hasil rekomendasi.

HOME  
(HTTPS://LIBRARY.GUNADARMA.AC.ID/DEPC  
SYSTEM/)

PEMBAYARAN  
(HTTPS://LIBRARY.GUNADARMA.AC.ID/DEPC  
SYSTEM/PEMBAYARAN)

BIODATA DIRI  
(HTTPS://LIBRARY.GUNADARMA.AC.ID/DEPC  
SYSTEM/ANGGOTA)

PEMINJAMAN BUKU  
(HTTPS://LIBRARY.GUNADARMA.AC.ID/DEPC  
SYSTEM/ANGGOTA/PEMINJAMAN)

PENULISAN ILMIAH  
(HTTPS://LIBRARY.GUNADARMA.AC.ID/DEPC  
SYSTEM/ANGGOTA/PENULISAN)

FOTOCOPY  
(HTTPS://LIBRARY.GUNADARMA.AC.ID/DEPC  
SYSTEM/ANGGOTA/FOTOCOPY)

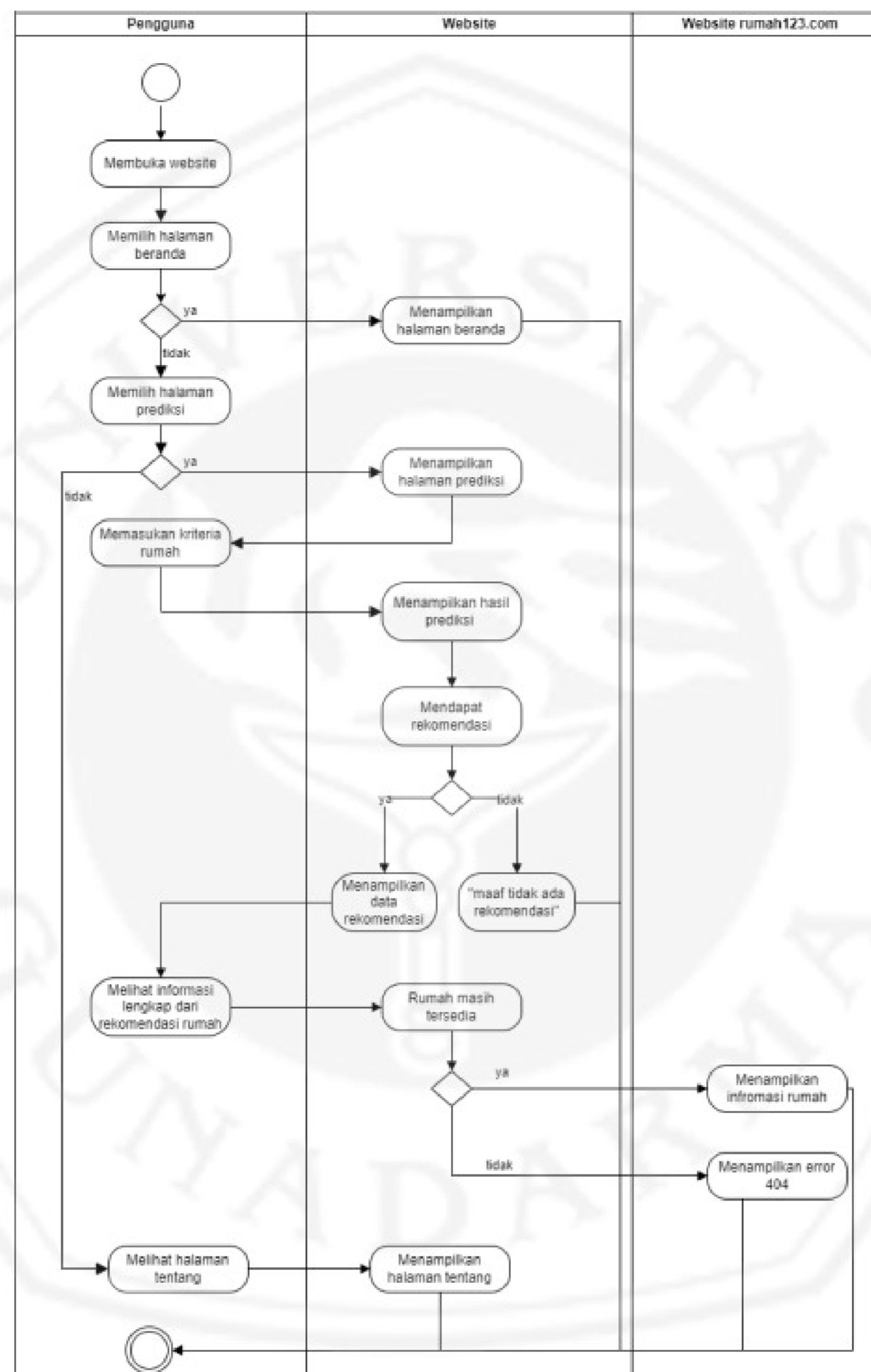
SUMBANG BUKU  
(HTTPS://LIBRARY.GUNADARMA.AC.ID/DEPC  
SYSTEM/ANGGOTA/SUMBANGBUKU)

KASUS & NOTIFIKASI  
(HTTPS://LIBRARY.GUNADARMA.AC.ID/DEPC  
SYSTEM/ANGGOTA/NOTIFIKASI)

BEBAS PERPUSTAKAAN  
(HTTPS://LIBRARY.GUNADARMA.AC.ID/DEPC  
SYSTEM/ANGGOTA/BEBASPERPUSTAKAAN)

KUNJUNGAN  
(HTTPS://LIBRARY.GUNADARMA.AC.ID/DEPC  
SYSTEM/KUNJUNGAN)

E-PAPER  
(HTTPS://LIBRARY.GUNADARMA.AC.ID/DEPC



Gambar 3. 22 Activity Diagram Website

- [!\[\]\(9e49868ba8eaf7071fc438017bcc77b8\_img.jpg\) HOME  
\(HTTPS://LIBRARY.GUNADARMA.AC.ID/DEPC  
SYSTEM/\)](#)
- [!\[\]\(88770677255ac3652019e5bd08cd9933\_img.jpg\) PEMBAYARAN  
\(HTTPS://LIBRARY.GUNADARMA.AC.ID/DEPC  
SYSTEM/PEMBAYARAN\)](#)
- [!\[\]\(92b47a52c91648e96ec79f56a96a075f\_img.jpg\) BIODATA DIRI  
\(HTTPS://LIBRARY.GUNADARMA.AC.ID/DEPC  
SYSTEM/ANGGOTA\)](#)
- [!\[\]\(cb1d67ac93897d8644fea0ec34acea7b\_img.jpg\) PEMINJAMAN BUKU  
\(HTTPS://LIBRARY.GUNADARMA.AC.ID/DEPC  
SYSTEM/ANGGOTA/PEMINJAMAN\)](#)
- [!\[\]\(050f024d470326ad37fb13c5634c31e9\_img.jpg\) PENULISAN ILMIAH  
\(HTTPS://LIBRARY.GUNADARMA.AC.ID/DEPC  
SYSTEM/ANGGOTA/PENULISAN\)](#)
- [!\[\]\(c61d5e36d2e1e315c2e1f558e9b3e0ce\_img.jpg\) FOTOCOPY  
\(HTTPS://LIBRARY.GUNADARMA.AC.ID/DEPC  
SYSTEM/ANGGOTA/FOTOCOPY\)](#)
- [!\[\]\(1112b4a375fc4fc31927149c9356181f\_img.jpg\) SUMBANG BUKU  
\(HTTPS://LIBRARY.GUNADARMA.AC.ID/DEPC  
SYSTEM/ANGGOTA/SUMBANGBUKU\)](#)
- [!\[\]\(868f51f3a1d817289e69e9a73d952326\_img.jpg\) KASUS & NOTIFIKASI  
\(HTTPS://LIBRARY.GUNADARMA.AC.ID/DEPC  
SYSTEM/ANGGOTA/NOTIFIKASI\)](#)
- [!\[\]\(091b6d4f9eba4aad245b668be9fd9320\_img.jpg\) BEBAS PERPUSTAKAAN  
\(HTTPS://LIBRARY.GUNADARMA.AC.ID/DEPC  
SYSTEM/ANGGOTA/BEBASPERPUSTAKAAN\)](#)
- [!\[\]\(76918a01f7427d2c19b445edda0dbbcf\_img.jpg\) KUNJUNGAN  
\(HTTPS://LIBRARY.GUNADARMA.AC.ID/DEPC  
SYSTEM/KUNJUNGAN\)](#)
- [!\[\]\(c439a2c02d95d389df48d83e059a6620\_img.jpg\) E-PAPER  
\(HTTPS://LIBRARY.GUNADARMA.AC.ID/DEPC](#)

Berdasarkan Gambar 3.22 di atas ada 3 *swim lane* sebagai proses bisnis dari aktifitas para pengguna yang menggunakan *website*. Aktifitas pertama yang akan dilakukan oleh pengguna adalah membuka halaman beranda *website* prediksi, lalu *website* akan menampilkan halaman beranda. Apabila pengguna tidak membuka halaman beranda dan membuka halaman prediksi maka *website* akan menampilkan halaman prediksi, lalu pengguna diharuskan untuk memasukkan kriteria rumah sehingga *website* akan menampilkan hasil prediksi. Jika data yang dimasukkan pengguna mendapatkan hasil rekomendasi maka *website* akan menampilkan data rekomendasi, namun apabila data yang dimasukkan tidak mendapatkan rekomendasi maka tidak akan menampilkan hasil rekomendasi. Apabila pengguna sudah melihat hasil prediksi dan rekomendasi maka pengguna juga dapat melihat informasi lengkap mengenai rumah rekomendasi yang diberikan jika rumah masih tersedia maka *website* rumah123.com akan menampilkan informasi lebih lengkap terkait rumah tersebut seperti detail dari lokasi rumah, gambar dari rumah yang ditawarkan dan informasi lainnya mengenai rumah yang ditawarkan, apabila rumah sudah tidak tersedia maka *website* rumah123.com akan menampilkan halaman *error 404* dikarenakan ada kemungkinan bahwa rumah sudah tidak di iklankan pada *website* rumah123.com. Pengguna juga dapat memilih halaman tentang maka *website* akan menampilkan halaman tentang.

### 3.6.1. Perancangan Tampilan Halaman Beranda

Beranda merupakan tampilan awal dari *website* prediksi yang menyampaikan beberapa informasi tentang rumah.

HOME  
(HTTPS://LIBRARY.GUNADARMA.AC.ID/DEPC  
SYSTEM/)

PEMBAYARAN  
(HTTPS://LIBRARY.GUNADARMA.AC.ID/DEPC  
SYSTEM/PEMBAYARAN)

BIODATA DIRI  
(HTTPS://LIBRARY.GUNADARMA.AC.ID/DEPC  
SYSTEM/ANGGOTA)

PEMINJAMAN BUKU  
(HTTPS://LIBRARY.GUNADARMA.AC.ID/DEPC  
SYSTEM/ANGGOTA/PEMINJAMAN)

PENULISAN ILMIAH  
(HTTPS://LIBRARY.GUNADARMA.AC.ID/DEPC  
SYSTEM/ANGGOTA/PENULISAN)

FOTOCOPY  
(HTTPS://LIBRARY.GUNADARMA.AC.ID/DEPC  
SYSTEM/ANGGOTA/FOTOCOPY)

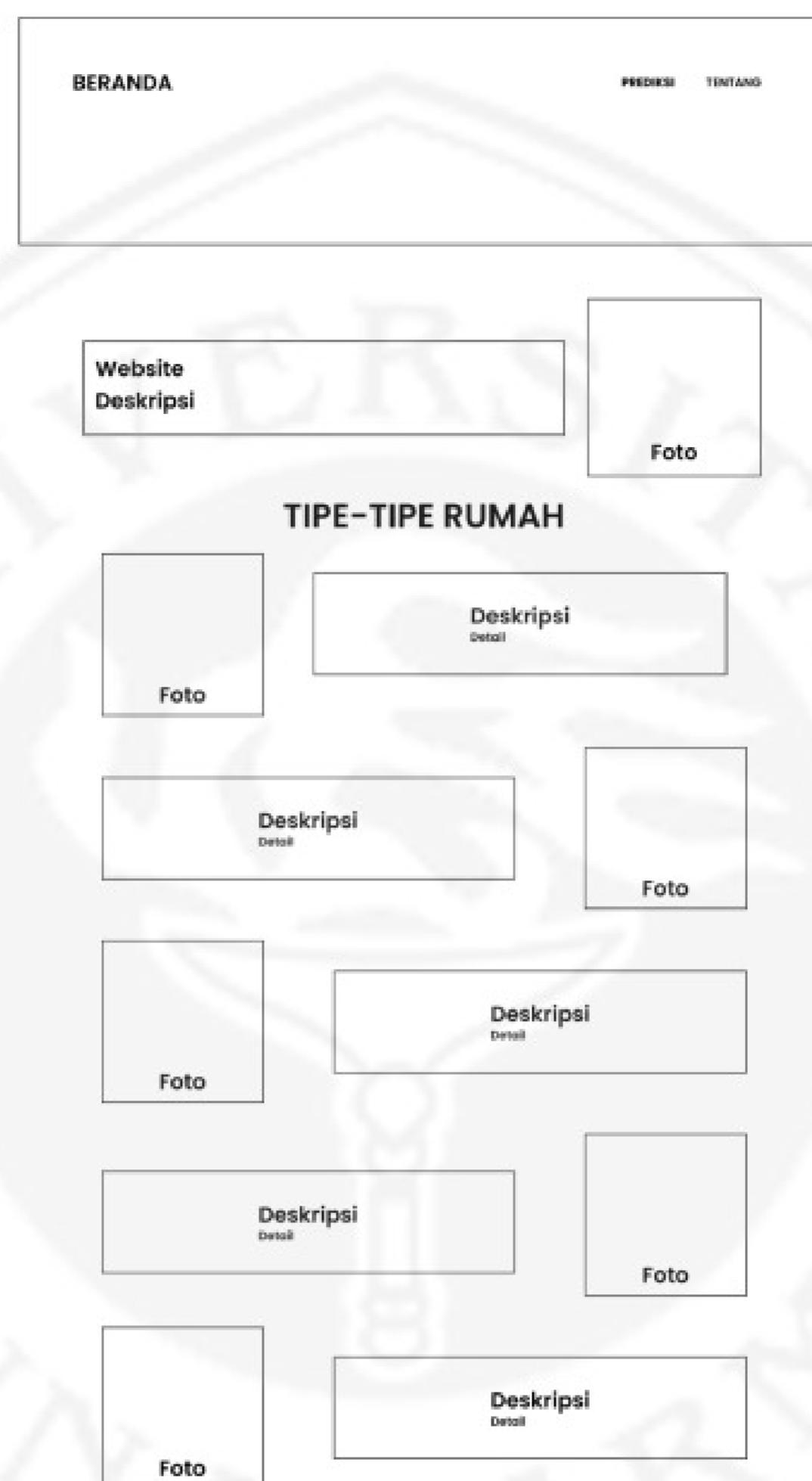
SUMBANG BUKU  
(HTTPS://LIBRARY.GUNADARMA.AC.ID/DEPC  
SYSTEM/ANGGOTA/SUMBANGBUKU)

KASUS & NOTIFIKASI  
(HTTPS://LIBRARY.GUNADARMA.AC.ID/DEPC  
SYSTEM/ANGGOTA/NOTIFIKASI)

BEBAS PERPUSTAKAAN  
(HTTPS://LIBRARY.GUNADARMA.AC.ID/DEPC  
SYSTEM/ANGGOTA/BEBASPERPUSTAKAAN)

KUNJUNGAN  
(HTTPS://LIBRARY.GUNADARMA.AC.ID/DEPC  
SYSTEM/KUNJUNGAN)

E-PAPER  
(HTTPS://LIBRARY.GUNADARMA.AC.ID/DEPC



Gambar 3. 23 Perancangan Tampilan Beranda

Berdasarkan Gambar 3.23 di atas merupakan perancangan dari tampilan beranda *website* dimana akan menampilkan tipe-tipe rumah dengan foto ilustrasi. Tipe-tipe rumah ini terdiri dari rumah 1 lantai, 2 lantai dan lainnya. Kemudian jika menekan detail akan masuk ke halaman detail tipe rumah.

HOME  
(HTTPS://LIBRARY.GUNADARMA.AC.ID/DEPC  
SYSTEM/)

PEMBAYARAN  
(HTTPS://LIBRARY.GUNADARMA.AC.ID/DEPC  
SYSTEM/PEMBAYARAN)

BIODATA DIRI  
(HTTPS://LIBRARY.GUNADARMA.AC.ID/DEPC  
SYSTEM/ANGGOTA)

PEMINJAMAN BUKU  
(HTTPS://LIBRARY.GUNADARMA.AC.ID/DEPC  
SYSTEM/ANGGOTA/PEMINJAMAN)

PENULISAN ILMIAH  
(HTTPS://LIBRARY.GUNADARMA.AC.ID/DEPC  
SYSTEM/ANGGOTA/PENULISAN)

FOTOCOPY  
(HTTPS://LIBRARY.GUNADARMA.AC.ID/DEPC  
SYSTEM/ANGGOTA/FOTOCOPY)

SUMBANG BUKU  
(HTTPS://LIBRARY.GUNADARMA.AC.ID/DEPC  
SYSTEM/ANGGOTA/SUMBANGBUKU)

KASUS & NOTIFIKASI  
(HTTPS://LIBRARY.GUNADARMA.AC.ID/DEPC  
SYSTEM/ANGGOTA/NOTIFIKASI)

BEBAS PERPUSTAKAAN  
(HTTPS://LIBRARY.GUNADARMA.AC.ID/DEPC  
SYSTEM/ANGGOTA/BEBASPERPUSTAKAAN)

KUNJUNGAN  
(HTTPS://LIBRARY.GUNADARMA.AC.ID/DEPC  
SYSTEM/KUNJUNGAN)

E-PAPER  
(HTTPS://LIBRARY.GUNADARMA.AC.ID/DEPC

### 3.6.2. Perancangan Tampilan Halaman Tentang

Halaman ini akan tampil jika menekan menu tentang.



Gambar 3. 24 Perancangan Tampilan Tentang

Berdasarkan Gambar 3.24 di atas tampilan tentang akan menunjukkan foto penulis dan nama dengan deskripsi. Terdapat informasi mengenai website prediksi ini yang akan menceritakan secara singkat kegunaan website beserta foto ilustrasi.

### 3.6.3. Perancangan Tampilan Halaman Prediksi

Pada halaman ini akan menampilkan kriteria inputan yang dibutuhkan untuk memproses prediksi harga dan memberikan rekomendasi rumah.

HOME  
(HTTPS://LIBRARY.GUNADARMA.AC.ID/DEPC  
SYSTEM/)

PEMBAYARAN  
(HTTPS://LIBRARY.GUNADARMA.AC.ID/DEPC  
SYSTEM/PEMBAYARAN)

BIODATA DIRI  
(HTTPS://LIBRARY.GUNADARMA.AC.ID/DEPC  
SYSTEM/ANGGOTA)

PEMINJAMAN BUKU  
(HTTPS://LIBRARY.GUNADARMA.AC.ID/DEPC  
SYSTEM/ANGGOTA/PEMINJAMAN)

PENULISAN ILMIAH  
(HTTPS://LIBRARY.GUNADARMA.AC.ID/DEPC  
SYSTEM/ANGGOTA/PENULISAN)

FOTOCOPY  
(HTTPS://LIBRARY.GUNADARMA.AC.ID/DEPC  
SYSTEM/ANGGOTA/FOTOCOPY)

SUMBANG BUKU  
(HTTPS://LIBRARY.GUNADARMA.AC.ID/DEPC  
SYSTEM/ANGGOTA/SUMBANGBUKU)

KASUS & NOTIFIKASI  
(HTTPS://LIBRARY.GUNADARMA.AC.ID/DEPC  
SYSTEM/ANGGOTA/NOTIFIKASI)

BEBAS PERPUSTAKAAN  
(HTTPS://LIBRARY.GUNADARMA.AC.ID/DEPC  
SYSTEM/ANGGOTA/BEBASPERPUSTAKAAN)

KUNJUNGAN  
(HTTPS://LIBRARY.GUNADARMA.AC.ID/DEPC  
SYSTEM/KUNJUNGAN)

E-PAPER  
(HTTPS://LIBRARY.GUNADARMA.AC.ID/DEPC

Gambar 3. 25 Perancangan Tampilan Prediksi

Berdasarkan Gambar 3.25 di atas tampilan prediksi ini akan meminta pengguna untuk memasukkan beberapa kriteria seperti pilih daerah, jumlah kamar tidur, dan lainnya. Tipe inputan pilih daerah akan dibuat secara dropdown, pada jumlah kamar tidur, kamar mandi dan garasi akan dibuat inputan numerik dengan rentang angka yang sudah ditentukan. Pada inputan luas tanah dan luas bangunan akan dibuat range input dengan rentang angka yang ditentukan lalu akan ada informasi mengenai rentang angka yang dipilih untuk luas tanah dan luas bangunan.

### 3.6.4. Perancangan Tampilan Halaman Hasil

Pada halaman ini akan tampil jika menekan tombol prediksi pada halaman prediksi.