

Pengembangan Sistem Prediksi Harga Pasar Properti Menggunakan *Big Data Platform*

Sinung Suakanto^{#1}, Aldi Christy^{*2}, Ventje J. L. Engel^{*3}, Dina Angela^{#4}

[#]Program Studi Teknik Elektro, Institut Teknologi Harapan Bangsa

Jalan Dipati Ukur 80 – 84 Bandung, Indonesia

¹sinung@ithb.ac.id

⁴dina_angela@ithb.ac.id

[#]Program Studi Sistem Komputer, Institut Teknologi Harapan Bangsa

Jalan Dipati Ukur 80 – 84 Bandung, Indonesia

²yonatanaldi@gmail.com

³ventje@ithb.ac.id

Abstract— *Property is an industry that will grow. Some of the problems in property sale and purchase transactions are lack of information about fair sale prices. This price information may be above or below market prices. This study will develop a model for predicting the market price of property in an area. The premise of this research is that the selling price of a house is proportional to the land area and building area in a particular area. The datasets used from this study come from information on selling prices and building area prices from several websites. The approach used in research is using regression and modification. The results of this study are expected to produce a model of recommended property market prices in an area.*

Keyword— *data analytics, property price, prediction, linear regression, NJOP.*

Abstrak— *Industri properti adalah salah satu industri yang diprediksi akan terus berkembang. Beberapa permasalahan yang muncul dalam transaksi jual-beli properti adalah sulitnya mendapatkan informasi mengenai harga jual properti yang wajar. Informasi harga tersebut bisa jadi berada di atas atau bawah harga pasaran. Penelitian ini akan mengembangkan model prediksi harga pasaran properti di suatu daerah. Premis dari penelitian ini adalah harga jual rumah sebanding dengan luas tanah dan luas bangunan di daerah tertentu. Dataset yang digunakan dari penelitian ini berasal dari informasi harga jual dan harga luas bangunan yang ada dari beberapa website. Pendekatan yang digunakan dalam penelitian adalah menggunakan regresi dan modifikasinya. Hasil dari penelitian ini diharapkan dapat menghasilkan sebuah model rekomendasi harga pasaran properti pada suatu daerah.*

Kata kunci— *analisis data, property, prediksi, regresi linear, NJOP.*

I. PENDAHULUAN

Industri properti di Indonesia adalah salah satu industri yang akan selalu ramai pembeli. Seiring berjalannya waktu penduduk di negara Indonesia akan terus bertambah dan tanah atau lokasi tempat tinggal yang ada akan terus berkurang. Hal ini menyebabkan permintaan terhadap properti, khususnya rumah tinggal, akan terus naik.

Pada saat seseorang ingin membeli sebuah properti terdapat masalah yang biasa ditemui, yaitu kurangnya informasi mengenai harga pasaran properti tersebut. Terbatasnya informasi mengenai harga tanah dan bangunan kemudian menyebabkan banyak masyarakat yang menjual atau membeli properti dengan buta informasi tentang harga pasaran.

Jika pembeli properti membeli dengan harga diatas pasaran, dampak yang paling terasa adalah pembeli mengeluarkan sumber daya lebih dari yang seharusnya. Sumber daya yang dikeluarkan pun biasanya tidak terbatas pada materi, tenaga manusia dan waktu pun masuk ke kategori tersebut.

Pada era digital dan informasi, sebenarnya data mengenai properti bisa diperoleh dari internet dengan lebih mudah. Data tersebut biasanya terdapat tersebar pada situs jual beli properti, situs resmi pemerintah, atau bahkan pada media sosial. Namun, tidak semua data yang ada di situs tersebut dapat memberikan gambaran umum tentang harga property di wilayah tertentu. Dengan memanfaatkan teknologi *big data*, data-data tersebut dapat diambil dan disimpan dengan skala dan jumlah yang besar berasal dari berbagai sumber.

Big data adalah sebuah sistem yang dapat menyimpan berbagai jenis data dengan kapasitas yang besar dan performa pengolahan data yang lebih cepat [1]. Selain menyimpan, *big data* pun dapat melakukan pemrosesan data dengan menggunakan fitur *big data analytic*. Pada kasus ini, implementasi *big data* dapat membantu menampung informasi properti, memproses informasi tersebut sesuai dengan kebutuhan masyarakat yang menggunakannya, dan memberikan rekomendasi langkah yang perlu diambil pengguna [2].

Hasil penelitian ini adalah sebuah system yang terdiri dari infrastruktur dan software analitik yang dapat membantu memberikan rekomendasi mengenai harga sebuah properti. Rekomendasi akan memberikan informasi apakah harga properti tersebut berada pada harga pasaran atau di luar harga pasaran pada lokasi tersebut. Rekomendasi tersebut akan diolah berdasarkan tiga buah data masukan, berupa harga

properti, luar tanah properti, dan luas bangunan properti tersebut. Pada penelitian ini, lokasi dibatasi untuk daerah Bandung Utara dan hanya pada kategori rumah dijual.

Diharapkan dengan adanya aplikasi yang dibangun dalam penelitian ini, pengguna dapat memperoleh rekomendasi untuk membantu proses transaksi (menjual dan membeli) property. Penelitian ini diharapkan juga dapat menjadi pondasi untuk mengembangkan system rekomendasi harga jual pasar properti.

II. METODOLOGI

Dalam penentuan harga properti, pemerintah hanya berperan sebagai penentu batas harga properti terendah. Batas tersebut diberi nama Nilai Jual Objek Pajak (NJOP)[2]. Untuk menentukan harga pasaran dari sebuah properti, masyarakat diberikan kebebasan dengan acuan NJOP yang secara umum kadang dijadikan sebagai patokan harga terendah. Harga properti sendiri akhirnya dapat dijual 1,5 – 2x dari harga NJOP [3]. Namun, kebanyakan masyarakat tidak mengetahui informasi mengenai NJOP tersebut sehingga mereka tidak mengetahui batas bawah dari harga pasaran dari property untuk wilayah atau daerah-daerah tertentu.

Masalah akan muncul ketika seseorang ingin membeli atau menjual sebuah properti. Jika ia akan membeli sebuah properti tanpa mengetahui informasi mengenai properti (harga tanah dan harga bangunan) tersebut terlebih dahulu, maka kemungkinan harga properti yang dibeli tersebut di atas harga pasaran. Jika seseorang akan menjual sebuah properti tanpa mengetahui informasi mengenai harga properti, maka kemungkinan ia akan menjual di bawah atau di atas harga pasaran. Jika penjual menjual di bawah harga pasaran, maka kemungkinan kerugiannya adalah waktu dan kesempatan yang terbuang karena properti tidak kunjung laku.

Dengan adanya *big data analytic*, data dari suatu properti dapat diolah untuk menghasilkan sebuah rekomendasi, misalnya, apakah harga properti tersebut berada di atas harga pasaran atau di bawah pasaran. *Big data analytic* sendiri adalah sebuah teknologi yang dapat digunakan untuk melakukan prediksi dengan data berukuran besar.

A. Riset Terkait

Banyak penelitian telah mengembangkan system rekomendasi atau system prediksi untuk berbagai keperluan. Misalnya untuk membantu memberikan rekomendasi untuk transaksi jual beli, pemilihan film maupun untuk media online.

1) Sistem Rekomendasi pada Jual Beli Online

Salah satu situs e-commerce besar seperti Amazon.com sudah mengimplementasikan konsep rekomendasi jual beli pada situsnya. *Big data* digunakan untuk menampung informasi pengguna situs tersebut, informasi transaksi yang dilakukan, dan preferensi dari tiap pengguna. Informasi tersebut kemudian dimasukkan ke dalam sistem rekomendasi situs Amazon.com [4].

Metode rekomendasi yang digunakan oleh Amazon.com pada situsnya adalah *Item-to-Item Collaborative Filtering* [5]

Pada Gambar 1 dijelaskan bagaimana metode tersebut bekerja. Metode ini bekerja dengan membandingkan barang yang ada di dalam keranjang belanja pengguna atau barang yang sudah dilihat pengguna dengan barang yang sudah dibeli dan yang dinilai oleh pengguna lainnya. Untuk menentukan barang apa yang akan direkomendasikan, maka akan dicari terlebih dahulu barang apa saja yang biasa dibeli dan barang yang akan dibeli pengguna yang sudah dinilai oleh pengguna tersebut. Hasil perbandingan tersebut kemudian disusun menjadi sebuah daftar rekomendasi [6].

2) Sistem Rekomendasi Film

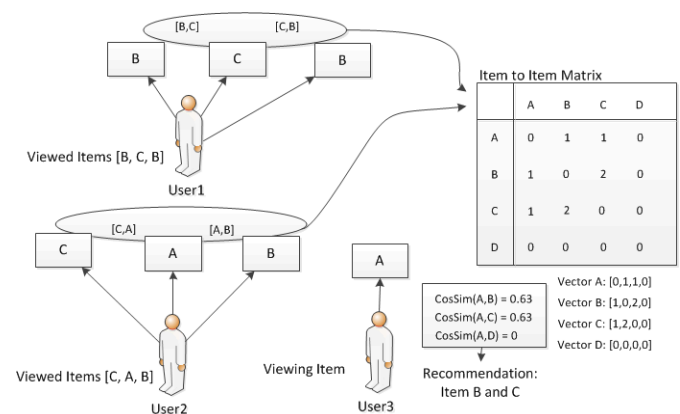
Pada tahun 2013, Netflix mengimplementasikan sistem *big data* dan *big data analytic* di dalam sistemnya. *Big data* digunakan untuk menampung data dari pengguna dan informasi preferensi dari pengguna tersebut. *Big data analytic* digunakan untuk melakukan rekomendasi film pada penggunaannya sesuai dengan preferensi yang dimilikinya [7]. Implementasi ini membuahkan hasil yang sangat memuaskan. Salah satu dampaknya adalah pertumbuhan jumlah pengguna layanan Netflix yang awalnya berjumlah 44,34 juta pengguna, pada tahun 2016 telah naik menjadi 93,80 juta pengguna [8].

3) Sistem Rekomendasi pada Media Sosial

Rekomendasi terkait penggunaan atau ketertarikan pada produk tertentu di media social juga telah dibangun. Pengguna layanan dapat memantau melacak, merekam, dan menganalisis setiap data yang ada di media sosial tersebut [9]. Data-data tersebut bisa berupa data mengenai akun media sosial, *hashtag*, atau topik yang sedang populer dibicarakan dalam media sosial, dan juga kata kunci pada mesin pencarian di dalam media sosial tersebut. Hasil dari proses tersebut dapat ditampilkan menggunakan grafik dan satuan tertentu.

B. Metode Solusi

Untuk menyelesaikan masalah terkait data yang ukurannya besar, diusulkan menggunakan sistem *big data* yang akan dibantu dengan perangkat *big data analytic*. Sistem akan menerima data mengenai properti yang didapat dari internet menggunakan teknik *web crawler*. Informasi yang diambil adalah harga, luas, dan lokasi dari properti yang dijual. Data tersebut kemudian akan diolah dan akan dibangun model pre-



Gambar 1 Cara kerja sistem rekomendasi Amazon.com

diksi terhadap harga pasaran properti dengan menggunakan model yang telah diusulkan sebelumnya [10],[11]. Hasil dari prediksi tersebut akan dicek kembali untuk melihat tingkat akurasinya.

C. Regresi Linear Sebagai Metode Prediksi

Salah satu teknik yang digunakan dalam data analitik adalah metode regresi linear. Metode regresi linear sendiri juga ada beberapa metode turunannya seperti metode regresi linear berganda. Regresi linear berganda digunakan untuk menguji hubungan linearitas antara beberapa variabel. Metode ini dipilih karena asumsi (premis) yang digunakan dalam penelitian, yaitu harga properti, akan dipengaruhi oleh beberapa faktor lainnya, seperti luas tanah, luas bangunan, dan sebagainya.

Pada beberapa studi kasus yang lain juga menggunakan metode regresi linear karena metode ini sangat cocok untuk melakukan prediksi [12],[13]. Pada penelitian tersebut ditunjukkan bahwa dengan menggunakan metode regresi linear untuk prediksi dengan *error rate* di kisaran 14% [12].

D. Analisis Permasalahan

Ada berbagai masalah terkait informasi penjualan properti yang sering muncul dan yang sering dihadapi para pelaku properti. Masalah-masalah tersebut terbagi dalam kategori sumber informasi properti, pengolahan informasi properti, dan penggunaan informasi properti. Penjelasan mengenai pembagian kategori ini dijelaskan dalam Tabel I.

Oleh karena adanya sejumlah keterbatasan dan kesulitan untuk memperoleh data transaksi jual-beli properti dan iklan pada media cetak, maka dalam penelitian ini akan digunakan data yang terdapat pada iklan media daring. Namun, data tersebut tentunya memiliki kekurangan. Kekurangan dari data yang diambil dari iklan media daring akan dijelaskan sebagai berikut:

1) Duplikasi iklan properti

Masalah yang muncul di dalam situs jual beli daring di Indonesia, yaitu adanya duplikasi iklan dari produk yang dijual. Sebuah barang yang sama dapat dijual oleh orang yang berbeda dan dapat dijual juga dengan harga yang berbeda. Pa-

da kasus ini banyak penjual properti yang melakukan duplikasi dari iklan properti yang sebenarnya untuk memenuhi *feed* dari halaman pencarian situs penjual properti.

2) Perbedaan informasi pada properti yang sama

Dikutip dari [10], informasi pemasaran yang menyesatkan menempati urutan ke-4 dari masalah transaksi perumahan di Indonesia. Informasi yang menyesatkan ini bisa dilakukan pihak pemasaran dengan berbagai cara. Cara-cara yang dilakukan biasanya dengan melakukan manipulasi pada iklan yang ditampilkan, melakukan promo penjualan properti tapi tidak menempatinya, menyantumkan nama dari daerah yang biasa dikenal di dalam iklan penjualan rumah, bahkan mengubah spesifikasi properti yang sebenarnya pada iklan.

3) Properti yang dijual sudah laku atau tidak dijual lagi

Dalam situs penjualan properti, tidak dicantumkan status sebuah properti sudah laku terjual atau tidak jadi dijual. Hal ini terjadi karena mayoritas situs penjualan properti tidak menyediakan fitur untuk mengganti status properti yang ditawarkan sudah terjual atau tidak lagi dijual. Faktor lain adalah para penjual properti biasanya tidak melakukan pembaharuan terhadap status properti. Pembeli properti harus menghubungi penjual yang memasang iklan properti untuk mengetahui apakah properti yang dipasarkan masih belum terjual atau sudah terjual.

E. Perancangan Sistem

Penelitian ini mengusulkan sebuah sistem dengan arsitektur yang terlihat pada Gambar 2. Arsitektur terdiri atas 3 buah modul, yaitu *Data Collecting*, *Data Analytic*, dan *Visualization*. Ketiga modul tersebut memiliki ketergantungan antara satu sama lain. Tugas dari masing-masing modul adalah mengumpulkan data, melakukan analisis terhadap data tersebut, dan menampilkan kesimpulan dari data yang telah diolah agar dapat dengan mudah dipahami penggunaanya.

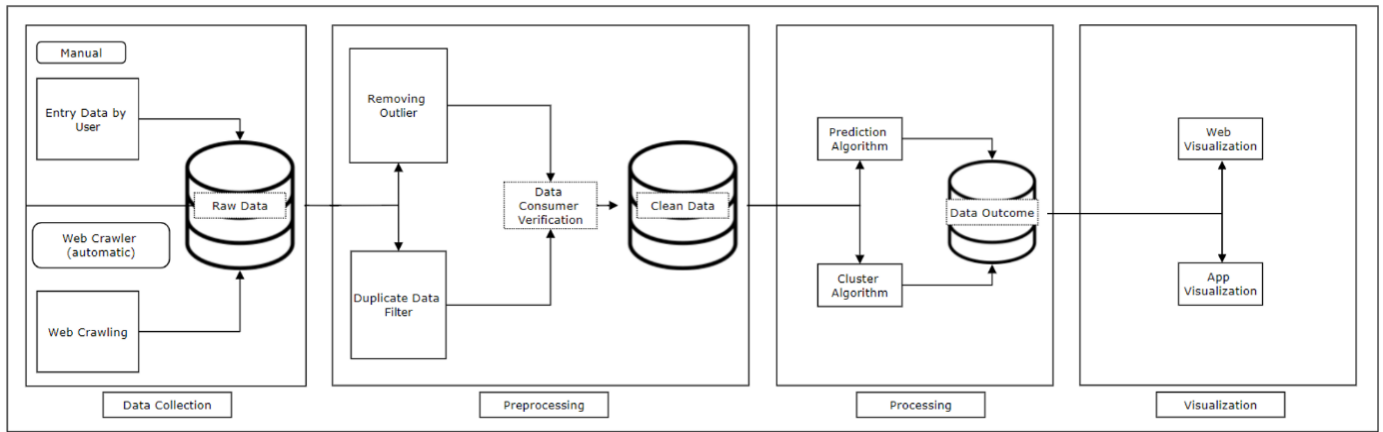
Data Collection dan *Preprocessing* merupakan proses pertama yang dilakukan. Data akan dikumpulkan dari situs penjualan properti di internet. Cara pengumpulan data pun akan dilakukan secara otomatis menggunakan sebuah program bernama *Crawler*. *Crawler* sendiri adalah sebuah program yang digunakan untuk mengambil data-data yang ada di internet secara otomatis.

Hasil dari modul *Data Collection* adalah sebuah basis data yang berisikan informasi harga-harga jual properti yang disimpan dalam basis data *big data*. Informasi tersebut mempunyai beberapa variabel, yaitu harga yang ditawarkan, luas properti, luas tanah properti, jumlah kamar tidur, dan jumlah kamar mandi. Variabel tersebut diambil karena diduga kuat memiliki dampak terhadap harga properti, menurut beberapa penelitian yang sudah dilakukan (lihat Tabel II). Pada Tabel II dijelaskan beberapa penelitian yang menggunakan variabel yang sama yang digunakan pada penelitian ini.

TABEL I

KATEGORI MASALAH DAN FOKUS PERMASALAHANNYA

No	Kategori Masalah	Fokus Permasalahannya
1.	Sumber Informasi	- Kebutuhan terhadap kuantitas informasi. - Kebutuhan terhadap kualitas informasi.
2.	Pengolahan Informasi	- Adanya kebutuhan pengolahan data. - Tidak adanya kemampuan pengolahan data.
3.	Penggunaan Informasi	- Membantu pengambilan keputusan penjual atau pembeli. - Membantu keputusan untuk mengambil investasi properti atau tidak.



Gambar 2 Arsitektur sistem yang diusulkan

TABEL II

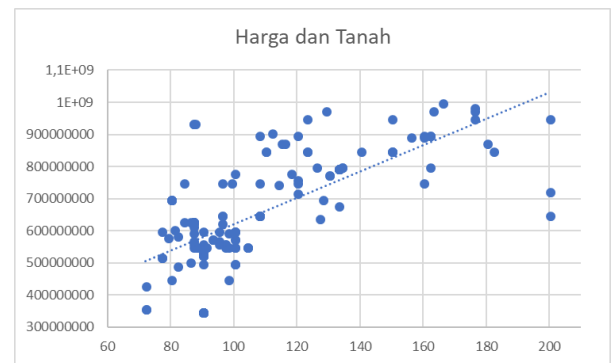
VARIABEL YANG DIGUNAKAN

No.	Variabel	Pustaka
1	Harga	[11], [13], [14], [15]
2	Luas Tanah	[11], [14], [15], [16]
3	Luas Bangunan	[11], [14], [15], [16]
4	Kamar Mandi	[11], [12], [14], [15]
5	Kamar Tidur	[11], [12], [14], [15]

Pada bagian *Processing*, data yang telah disaring akan diterima akan diolah dengan mempertimbangkan variabel yang telah ditetapkan. Variabel tersebut terdiri atas nama penjual, lokasi properti dijual, harga properti, luas tanah properti, luas bangunan properti, nomor telepon penjual, jumlah kamar, jumlah kamar mandi, dan alamat situs properti. Dari seluruh variabel tersebut akan digunakan lima variabel untuk menghasilkan analisis berdasarkan daerah dari properti yang dijual, yaitu harga properti, luas tanah, luas bangunan, jumlah kamar tidur dan jumlah kamar mandi.

Setelah menerima data mentah tersebut, sistem akan melakukan penyaringan terhadap data. Terdapat dua metode penyaringan yang dilakukan terhadap data mentah tersebut. Metode penyaringan pertama adalah melakukan penyaringan terhadap data pencilan atau data yang tidak valid, atau *removing outlier*, dan metode yang kedua adalah melakukan penyaringan terhadap data yang memiliki duplikasi, atau *duplicate data filter*.

Membuang data pencilan (*outlier*) dilakukan dengan melihat apakah data yang dimiliki memiliki keterkaitan yang erat satu dengan yang lainnya. Hal ini bisa dilihat dengan melakukan pemetaan terhadap data yang dimiliki dan melihat keterkaitan dari variabel yang dimiliki data tersebut dengan variabel lainnya. Pada Gambar 3 terlihat hubungan antara variabel harga jual dan luas tanah, terlihat juga pada gambar garis *trendline* yang menunjukan keterkaitan kedua variabel tersebut. Semakin besar luas tanah yang dimiliki, maka semakin tinggi harga jual yang ditawarkan. Jika data yang dimiliki berada jauh dari nilai rata-rata garis *trendline*, maka data akan dihapus.



Gambar 3 Hubungan variabel harga jual dan luas tanah

Untuk melakukan penghapusan terhadap data duplikat, data yang dimiliki akan diurutkan menurut variabel yang dimiliki. Dengan melakukan pengurutan tersebut dapat terlihat data yang memiliki duplikat. Data yang memiliki kesamaan dengan data yang lainnya akan dihapus. Data yang tidak memiliki kesamaan atau unik akan disimpan dan masuk kedalam tahap penyesuaian data selanjutnya.

Setelah data melewati dua tahap tersebut, selanjutnya data akan diverifikasi terlebih dahulu sebelum digunakan pada tahap selanjutnya. Hal ini dilakukan untuk membantu meningkatkan akurasi algoritme rekomendasi yang akan digunakan. Pada tahap verifikasi, data akan diformat ulang dengan tujuan untuk mempermudah sistem melakukan analisis.

Sistem juga akan memiliki beberapa ketentuan terhadap variabel yang digunakan. Variabel harus merupakan angka me memudahkan sistem melakukan analisis. Jika variable memiliki karakter atau bahkan tanda baca, maka sistem akan menghapus baris data tersebut.

Dengan menggunakan lima buah variabel yang telah diambil, langkah selanjutnya adalah melakukan analisis berdasarkan lokasi dari properti yang dijual. Analisis yang dimaksud bertujuan untuk melihat apakah properti yang ditawarkan tersebut memiliki nilai yang wajar atau tidak. Untuk melihat apakah wajar atau tidak, analisis dari sistem akan menghasilkan estimasi harga tanah dan harga bangunan per meter di daerah tersebut. Dari hasil estimasi itu kemudian

dibandingkan dengan harga properti pasaran lalu rekomendasi diberikan kepada pengguna.

Pada sistem akan digunakan algoritme untuk melakukan analisis data yang didasari oleh metode regresi linear berganda. Alasan menggunakan regresi linear berganda adalah karena metode dan rumus yang dihasilkan dari analisis dapat dimasukkan ke dalam model penentu harga properti. Metode regresi linear berganda mempunyai dua jenis variabel, yaitu variabel bebas dan variabel terikat. Pada penelitian ini variabel terikat yang dimiliki adalah harga dari properti yang dijual, sedangkan variabel yang terikat adalah luas properti, luas tanah, jumlah kamar, dan jumlah kamar mandi.

Setelah *dataset* terkumpul maka dibuat sebuah model prediksi untuk menentukan harga prediksi pasar yang ditentukan oleh variabel-variabel tersebut. Secara umum, model prediksi didekati dengan formula sebagai berikut:

$$H_{pd} = f(H_B, H_T) = \left((H_{TPM} \times L_T) + (H_{TBM} \times L_B) \right) + K_d \quad (1)$$

H_{pd} adalah harga penawaran pada sebuah daerah, H_T adalah harga tanah yang didapat dari $(HTPM \times L_T)$, H_B adalah harga bangunan yang didapat dari $(HTBM \times L_T)$, K_d adalah konstanta daerah, H_{TPM} adalah koefisien harga tanah per meter, L_T adalah luas tanah, H_{BPM} adalah koefisien harga bangunan per meter, L_B adalah luas bangunan.

Penyelesaian dari model ini dapat menggunakan metode regresi linear berganda, di mana rumus dasar regresi linear berganda yang digunakan adalah sebagai berikut:

$$y = a + (b(1)X(1)) + (b(2)X(2)) \quad (2)$$

a adalah K_d , $b(1)$ dan $b(2)$ masing-masing adalah HTPM dan HBPM, dan $X(1)$ dan $X(2)$ masing-masing adalah LT (luas tanah) dan LB (luas bangunan).

III. HASIL DAN PEMBAHASAN

A. Implementasi Perangkat Lunak

Perangkat *big data* yang akan digunakan adalah Apache Hadoop dan Apache Spark. Adapun *tools* untuk data analitiknya menggunakan R. Fungsi dari masing-masing perangkat tersebut adalah penyimpanan dan pengolahan data. *Dataset* diambil untuk data-data harga jual rumah dan tanah sebanyak 7.241 data yang ada di wilayah Bandung Utara, Bale Endah untuk Bandung Selatan, Padalarang untuk Bandung Barat, dan Arcamanik untuk Bandung Timur. Data-data ini akan dibersihkan dahulu untuk mendapatkan hasil yang lebih baik dalam proses pengembangan model prediksinya.

Sebelum data diolah, data harus dibersihkan terlebih dahulu. Yang dimaksudkan dengan pembersihan adalah pembuangan data yang tidak sesuai dengan format yang ditentukan dan data yang tidak terambil oleh *Crawler* atau kosong. Data yang diambil oleh *Crawler* dipisahkan menurut daerah dari kota yang akan diolah. Dalam penelitian ini kota

yang diambil adalah Bandung dan dibagi menjadi Bandung Utara, Bandung Selatan, Bandung Timur, dan Bandung Barat. Setelah itu, tiap data dari tiap daerah akan dibersihkan, setiap baris yang memiliki variabel yang kosong akan dihapus dan disesuaikan dengan format untuk pengolahan data. Hal ini dapat terjadi karena adanya kesalahan dalam mengambil data oleh *Crawler*. Kesalahan terjadi karena format pengambilan data yang digunakan oleh *Crawler* tidak sesuai dengan format data yang ada pada situs. Format yang dimaksud adalah format untuk tiap variabel atau fitur yang dimiliki oleh daerah tersebut. Dalam penelitian ini, ada 5 fitur yang diperlukan untuk melakukan pengolahan data. Fitur tersebut adalah, harga properti, luas tanah properti, luas bangunan properti, jumlah kamar tidur, dan jumlah kamar mandi. Semua fitur tersebut memiliki jenis tipe *integer*, sehingga jika fitur memiliki karakter maka baris dari data akan dihapus.

Variabel lokasi digunakan untuk memisahkan data berdasarkan daerah, dan *desc* adalah variabel yang digunakan untuk membantu menunjukkan duplikasi data dalam sebuah daerah. Pemisahan data berdasarkan daerah dilakukan karena setiap daerah memiliki faktor penentu harga yang berbeda daerah satu dengan yang lainnya.

Setiap variabel dilihat korelasinya terhadap harga jual. Hal ini dilakukan untuk melihat variabel apa saja yang akan mempengaruhi harga dari sebuah properti. Gambar 4 menjelaskan hubungan masing-masing variabel dengan harga. Dari gambar tersebut dapat dilihat bahwa luas tanah dan luas bangunan mempunyai korelasi yang kuat terhadap harga jual sebuah properti. Hal ini ditunjukkan dengan semakin tinggi luas tanah atau luas bangunan, maka harga jual properti tersebut juga akan semakin tinggi. Lain halnya dengan jumlah kamar dan kamar mandi yang berkebalikan. Jumlah kamar tidur dan kamar mandi tidak memiliki pengaruh yang signifikan terhadap harga jual properti. Dari kelima variabel yang dikumpulkan, banyaknya kamar tidur dan kamar mandi terlihat tidak memberikan pengaruh terhadap harga jual sebuah properti. Hal ini membuat pengolahan data tidak perlu menggunakan kedua variabel tersebut dan membuat analisis dapat dilakukan dengan menggunakan tiga buah variabel saja, yaitu harga, luas tanah, dan luas bangunan properti.

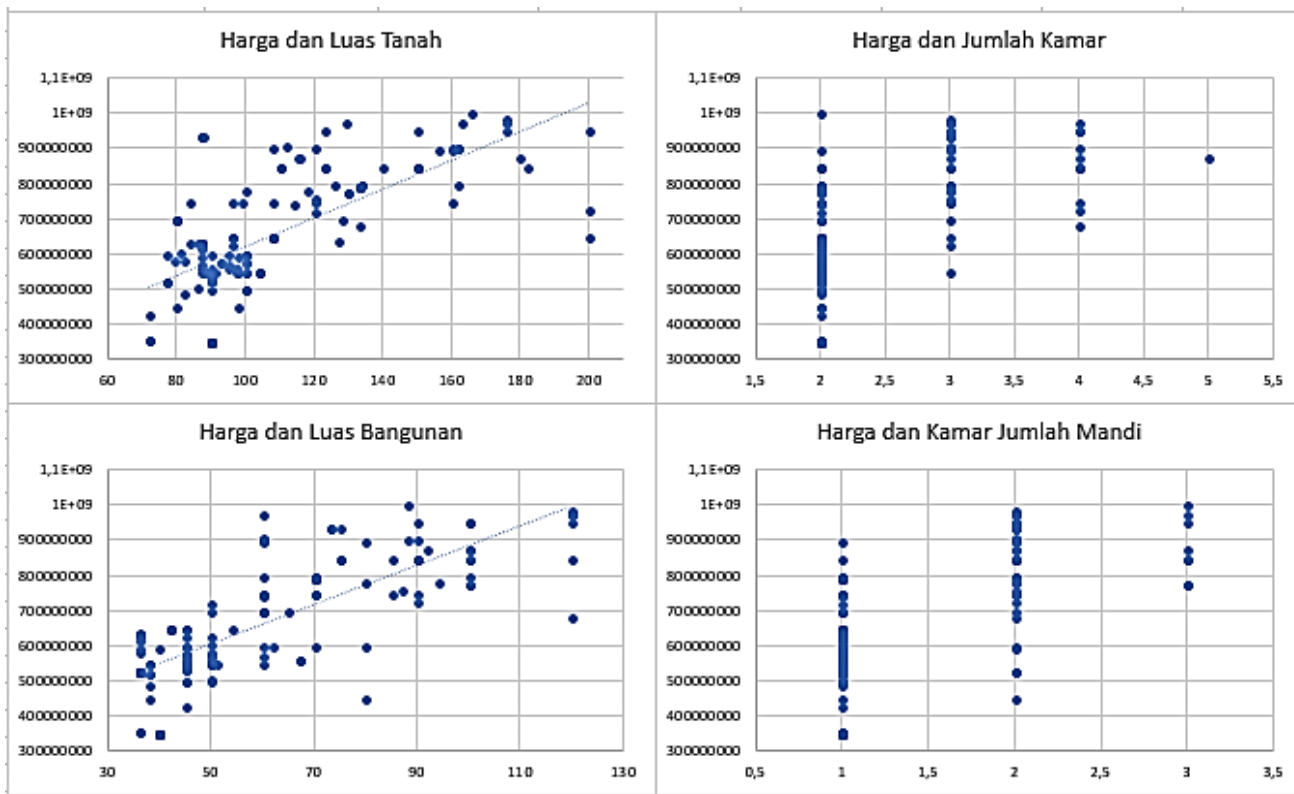
B. Hasil Pengolahan

Hasil pengolahan dan analisis data properti dari empat daerah di kota Bandung dijabarkan dalam bagian ini. Metode pengolahan yang digunakan metode *Double Regression* (DR). Analisis ini menggunakan metode regresi linear berganda sebanyak dua kali. Hasil analisis berupa prediksi harga bangunan dan harga tanah per meter. Hasil prediksi tersebut kemudian dikalikan dengan luas tanah dan luas bangunan dan bangunan dan harga tanah per meter. Hasil prediksi tersebut kemudian dibandingkan dengan harga properti yang ditawarkan. Setelah dibandingkan, akan terlihat tingkat kesalahan (ada selisih harga) pada setiap properti yang ditawarkan (*error rate*) seperti yang terlihat pada Tabel III.

Tabel III menunjukan hasil dari metode analisis DR. Dengan metode DR ini dapat dibangun model untuk menentukan pre-

diksi harga pasaran dengan menggunakan variabel luas tanah dan luas bangunan. Secara umum, terlihat bahwa untuk daerah/wilayah yang berbeda, menghasilkan model yang berbeda. Hasil prediksi menunjukkan hasil koefisien harga ta-

nah, koefisien harga bangunan, dan konstanta yang berbeda. Terjadi perbedaan nilai koefisien untuk setiap daerah yang berbeda. Dengan demikian, model prediksi untuk setiap wilayah daerah pun akan berbeda-beda.



Gambar 2 Gambar sebaran data dengan harga

TABEL III
HASIL METODE ANALISIS DR

Daerah		Jumlah Data	Metode 3 (DR)			Error Rate
			Koefisien Harga Tanah	Koefisien Harga Bangunan	Konstanta	
Utara	Setiabudi	12	2.225.533,66	2.150.233,41	226.946.608,36	0,12
Selatan	Baleendah	22	6.236.452,19	(755.816,19)	(131.006.613,29)	13,80
Barat	Padalarang	12	7.845.951,98	1.440.249,09	(240.233.718,17)	14,32
Timur	Arcamanik	19	5.261.001,57	518.816,43	16.133.744,73	2,14

IV. KESIMPULAN

Penelitian ini menunjukkan bahwa pemrosesan data-data properti dapat menggunakan teknologi *big data* dan *data analitik*. Model prediksi yang dikembangkan telah dapat menghasilkan model prediksi untuk setiap wilayah daerah. Setiap wilayah daerah ternyata memiliki model prediksi yang berbeda-beda sesuai dengan karakteristik daerah tersebut. Permasalahan dari penelitian ini adalah data-data yang diperoleh setelah data dibersihkan menjadi sangat sedikit sehingga model prediksinya menjadi kurang akurasi. Pada penelitian ke depan dapat dikembangkan juga *dataset* untuk tahun yang berbeda untuk melihat konsistensi setiap tahunnya.

Penelitian ini masih banyak ruang untuk pengembangan lebih lanjut, misalnya dengan pemanfaatan teknologi Cloud.

DAFTAR REFERENSI

- [1] F. Azkia (2016, Februari 15). *Apa itu NJOP dan NJKP* [Online]. Tersedia: <https://www.rumah.com/berita-properti/2016/2/117447/apa-itu-njop-dan-njpk>.
- [2] A. Bhatia, *Big Data Analytics*, MU CBGS, 2016.
- [3] W. Brand, *Big Data for Dummies*, First Edition, New Jersey: John Wiley & Sons, Inc, 2013.
- [4] W. Ardiyanto (2017, Mei 21). *Mau Survei Properti? Pahami Peran Semua Pihak Yang Terlibat!* [Online]. Tersedia: <https://www.rumah.com/berita-properti/2017/5/152675/mau-survei-properti-pahami-peran-semua-pihak-yang-terlibat>.
- [5] T. Kwaiec. *The Amazon Recommendations Secret to Selling More Online* [Online]. Available: <http://rejoiner.com/resources/amazon-recommendations-secret-selling-online/>
- [6] G. Linden, B. Smith, J. York, "Amazon.com Recommendations: Item-to-Item Collaborative Filtering", *IEEE Internet Computing*, Vol. 7, pp. 76-80, January 2003.
- [7] B. Marr (2015, April 22). *Big Data: How Netflix Uses It to Drive Business Success* [Online]. Available: <https://www.smartdatacollective.com/big-data-how-netflix-uses-it-drive-business-success/>.
- [8] Z. Bulygo. *How Netflix Uses Analytics To Select Movies, Create Content, and Make Multimillion Dollar Decisions* [Online]. Available: <https://blog.kissmetrics.com/how-netflix-uses-analytics/>
- [9] I. Kurniawan, *Alasan Mengapa Transaksi Jual Beli Properti Belum Bisa Online*, [Online]. Tersedia: <https://id.techinasia.com/alasan-mengapa-transaksi-jual-beli-properti-belum-bisa-online>
- [10] Y. Putri, *Segudang Masalah Konsumen Perumahan* [Online], Tersedia: <http://ylki.or.id/2011/10/segudang-masalah-konsumen-perumahan/>
- [11] J. Yang, "Housing Price Prediction Using Support Vector Regression," *Master's Projects*, Mei 2017.
- [12] V. Hugo, "Property valuation using machine learning algorithms a study in a metropolitan - area of Chile," *Conference: AMSE Conference Santiago/Chile*, Januari 2016.
- [13] Aaron Ng, "Machine learning for a London housing price prediction mobile application," Final Project, Department of Computing, Imperial College London, Juni 2015.
- [14] A. Seutin, *Using Machine Learning to Predict Housing Price Given Multivariate Input*, 2016.
- [15] V. Venkat, S. Vijay, S. Banu, "Identifying Customer Interest in Real Estate Using Data Mining Techniques," *International Journal of Computer Science and Information Technologies*, Vol. 5, No. 3, pp. 3081-3084, 2014.
- [16] A. Nur, R. Ema, H. Taufiz, W. Firdaus, "Modeling House Price Prediction using Regression Analysis and Particle Swarm Optimization (Case Study: Malang, Indonesia)," *International Journal of Advanced Computer Science and Applications*, Vol. 8, No. 10, pp. 323-326, 2017.

Sinung Suakanto, kelahiran Klaten tahun 1982 dan memperoleh gelar Sarjana Teknik di bidang Teknik Elektro. Melanjutkan pendidikan doctoral di bidang Teknik Elektro dalam bidang jaringan komunikasi di ITB. Minat penelitian pada bidang jaringan sensor, cloud computing, serta teknologi informasi. Saat ini aktif sebagai staf pengajar di Teknik Elektro dan Sistem Komputer ITHB.

Aldi Christy, menerima gelar Sarjana Komputer dari Institut Teknologi Harapan Bangsa jurusan *Media and Internet Technology* pada tahun 2018.

Ventje Jeremias Lewi Engel, memperoleh gelar Sarjana Teknik dari Sistem dan Teknologi Informasi ITB pada Oktober 2012 dan Magister Teknik dari Informatika ITB pada Juli 2013. Minat penelitian pada bidang analisis data, *Internet of Things*, dan keamanan informasi. Saat ini aktif sebagai staf pengajar di Prodi Sistem Komputer ITHB, Bandung.

Dina Angela, menyelesaikan S1 Jurusan Teknik Elektro bidang Telekomunikasi di Universitas Kristen Maranatha pada 1999 dan S2 Jurusan Teknik Elektro bidang Telekomunikasi di ITB pada 2003. Bidang penelitian: antena dan propagasi dan sistem komunikasi.

Halaman kosong