

Algoritma *K-Nearest Neighbour* untuk Memprediksi Harga Jual Tanah

Wiyli Yustanti *

Abstrak

Sampai saat ini masih banyak masyarakat yang kesulitan untuk menentukan pilihan dalam memilih tanah yang strategis dengan harga sesuai kemampuan karena kurangnya pengetahuan tentang harga tanah berdasarkan harga pasar. Berdasarkan permasalahan tersebut maka dilakukan perancangan dan pembuatan aplikasi yang dapat digunakan untuk memprediksi harga jual tanah dengan pendekatan algoritma *K-Nearest Neighbour* (KNN). Dengan aplikasi ini diharapkan dapat memberikan Informasi yang lebih akurat dan efisien tentang harga jual tanah serta membantu para calon pembeli atau penjual tanah untuk memprediksi nilai tanah sesuai dengan kriteria yang ditentukan. Data yang dikumpulkan berupa data sekunder. Metode yang digunakan adalah gabungan antara tahapan data mining yang dikenal dengan istilah *Cross-Industry Standard Process for Data Mining* (CRISP-DM) dan metode pengembangan perangkat lunak *Waterfall Model*. Secara keseluruhan aplikasi ini mampu untuk memprediksi nilai tanah dengan pemrosesan yang cukup lama karena algoritma KNN prinsipnya adalah membandingkan data *testing* (data baru) dengan data *training* (data lama) secara satu persatu. Hasil akurasi dari prediksi data testing adalah sebesar 80%.

Kata Kunci: Prediksi, harga tanah, *K-Nearest Neighbour*.

Abstract

Until now there are many people who gets difficulties to make a choice in selecting strategic land because of lack of knowledge about the land price based on market prices. Of the problem then do the design and creates an application that can be used to predict the selling price of land with the approach of *K-Nearest Neighbor* algorithm (KNN). With this application is expected to provide more accurate information and efficient regarding the selling price of land and help potential buyers or sellers of land to predict the value of the land in accordance with the specified criteria. Data collected in the form of secondary data. The method used is a combination of data mining stage known as *Cross-Industry Standard Process for Data Mining* (CRISP-DM) and the method of *Waterfall Model* of software engineering.

Overall, this application is able to predict the value of land with a fairly long process because the principle of the KNN algorithm is to compare testing data (new data) with training data (old data) in one by one. The results of the prediction accuracy of testing data is about 80%.

Keywords: *Prediction, Land Price, K-Nearest Neighbour*.

* Prodi Manajemen Informatika, Jurusan Teknik Elektro, Univesitas Negeri Surabaya, email: wiyli5@gmail.com.

1. Pendahuluan

Pertumbuhan penduduk yang semakin meningkat mengakibatkan permintaan terhadap tanah juga semakin tinggi. Tentu saja bukan hanya angka pertumbuhan atau tingkat populasi yang harus diperhitungkan, akan tetapi keberadaan akan dana yang dimiliki masyarakat juga sangat mempengaruhi. Ditambah lagi pertumbuhan ekonomi sebuah negara, juga ikut mempengaruhi pergerakan harga tanah. Apabila tanah yang tersedia tidak dapat memenuhi semua permintaan maka akibat selanjutnya adalah kenaikan harga tanah. Hal ini yang menyebabkan trend harga tanah semakin meningkat dari tahun ke tahun.

Kekuatan penawaran dan permintaan (*supply and demand*) saling berinteraksi mempengaruhi nilai tanah yang direfleksikan oleh harga penjualan. Dalam jangka pendek, penawaran menjadi sangat kaku (*inelastic*), karena luas tanah tidak dapat ditambah secara cepat dan drastis. Sementara itu kebutuhan akan tanah sebagai tempat tinggal atau tempat usaha maupun sebagai barang investasi semakin lama semakin mendekati gejala konsumtif (*durable consumption goods*).

Pada dasarnya harga tanah ditentukan oleh dua faktor yaitu faktor terukur (*tangible factor*) seperti lokasi tanah yang strategis, dekat dengan pusat kota dan faktor tidak terukur (*intangibile factor*). Faktor tidak terukur lebih banyak disebabkan oleh aspek sosio-psikologis sehingga sulit untuk dikaji secara ilmiah, misalnya kenyamanan dan keamanan.

Sampai saat ini masih banyak masyarakat yang kesulitan untuk menentukan pilihan dalam memilih tanah yang strategis dengan harga sesuai kemampuan karena kurangnya pengetahuan tentang harga tanah berdasarkan harga pasar. Permasalahan yang akan diselesaikan dalam penelitian ini adalah bagaimana memprediksi harga jual tanah berdasarkan faktor-faktor yang dapat terukur. Untuk menyelesaikannya akan dilakukan percobaan berdasarkan data-data yang dapat diukur seperti jarak lokasi dengan pusat kota, sekolah, rumah sakit dan sebagainya. Pendekatan yang digunakan untuk memprediksi harga jual tanah adalah algoritma *K-Nearest Neighbour* (KNN).

2. Tinjauan Pustaka

Terkait dengan permasalahan yang akan diselesaikan, kajian pustaka yang penting untuk dipahami adalah teori mengenai penentuan harga jual tanah dan metode *K-Nearest Neighbour*.

2.1. Nilai Tanah

Tanah adalah salah satu faktor produksi yang tidak bisa diproduksi oleh manusia. Pemanfaatan tanah yang baik akan menjamin kelangsungan ekosistem yang stabil, membatasi pencemaran udara, serta dapat menciptakan struktur politik, ekonomi, sosial, budaya, pertahanan, keamanan nasional masyarakat [7].

Ray M. Northam (1975), lihat [4], mengemukakan dua buah pengertian tentang nilai tanah, yakni:

1. Nilai tanah adalah nilai pasar (*market value*) yaitu harga jual beli tanah yang terjadi pada suatu waktu tertentu.
2. Nilai tanah adalah nilai assessment (*assessed value*) yaitu nilai yang diestimasi oleh seorang penilai. *Market value* merupakan data dasar bagi *assessed value*.

Menurut Shenkel, seperti yang dikutip oleh [8], nilai tanah mempunyai definisi atau pengertian bermacam-macam tergantung pada konteks dan tujuannya serta sudut pandangnya.

Nilai tanah secara definisi diartikan sebagai kekuatan nilai dari tanah untuk dipertukarkan dengan barang lain. Sebagai contoh tanah yang mempunyai produktivitas rendah seperti tanah padang rumput relatif lebih rendah nilainya karena keterbatasan dalam penggunaannya. Sedangkan nilai pasar tanah didefinisikan sebagai harga (yang diukur dalam satuan uang) yang dikehendaki oleh penjual dan pembeli.

Nilai tanah dalam konteks pasar properti adalah nilai pasar wajar yaitu nilai yang ditentukan atau ditetapkan oleh pembeli yang ingin membeli sesuatu dan penjual ingin menjual sesuatu berdasarkan persetujuan atau kesepakatan kedua belah pihak dalam kondisi wajar tanpa ada tekanan dari pihak luar pada proses transaksi jual beli sehingga terjadi kemufakatan. Pembeli dan penjual mempunyai tenggang waktu yang cukup atas properti yang diperjualbelikan dan bertindak untuk kepentingan sendiri.

Menurut Peace and Turner yang dikutip oleh [8], faktor non-manusia berkenaan dengan eksternalitas yang diterima oleh tanah tersebut. Jika eksternalitas bersifat positif, seperti dekat dengan pusat perekonomian, bebas banjir, kepadatan penduduk, dan adanya sarana jalan, maka tanah akan bernilai tinggi jika dibandingkan dengan tanah yang tidak menerima eksternalitas, meskipun luas dan bentuk tanah itu sama. Jika tanah menerima eksternalitas yang bersifat negatif, seperti dekat dengan sampah, jauh dari pusat kota/perekonomian, tidak bebas banjir, maka tanah akan bernilai rendah jika dibandingkan dengan tanah yang tidak menerima eksternalitas yang negatif [5].

Untuk melakukan penilaian tanah, perlu diketahui beberapa prinsip penilaian. Menurut Joseph K. Eckert, seperti yang dikutip oleh [8], mengemukakan empat prinsip penilaian tanah, yakni penawaran dan permintaan (*supply and demand*), penggunaan yang tertinggi dan terbaik (*highest and the best use*), keuntungan produktivitas (*surplus productivity*), serta prinsip perubahan danantisipasi (*change and anticipation*).

2.2 Faktor – Faktor yang Mempengaruhi Nilai Tanah

Dalam jurnal *American Institute of Real Estate Appraisers* [10], ada empat faktor yang mempengaruhi nilai tanah, yaitu

1. Faktor Ekonomi

Faktor ekonomi berkaitan dengan keadaan ekonomi global/internasional, nasional, regional maupun lokal. Variabel-variabel permintaan (*demand*) yang mempengaruhi nilai tanah termasuk di dalamnya ialah jumlah tenaga kerja, tingkat upah, tingkat pendapatan dan daya beli, tersedianya keuangan, tingkat suku bunga dan biaya transaksi.

2. Faktor Sosial

Faktor sosial membentuk pola penggunaan tanah pada suatu wilayah. Kepadatan penduduk, tingkat pendidikan, tingkat kejahatan dan kebanggaan memiliki (daerah bergengsi) adalah faktor-faktor sosial yang mempengaruhi nilai tanah.

3. Faktor Politik dan Kebijakan Pemerintah

Kebijakan pemerintah di bidang hukum dan politik mempengaruhi nilai tanah. Beberapa contoh kebijakan yang dapat mempengaruhi biaya dan alokasi penggunaan tanah yang pada gilirannya akan meningkatkan harga tanah, antara lain; kebijakan pemilikan sertifikat tanah, peraturan penataan ruang dengan penentuan mintakat atau zoning, peraturan perpajakan, peraturan perijinan (SIPPT, IMB dan lain-lain) ataupun penentuan tempat pelayanan umum (sekolah dasar, rumah sakit, dan lain-lain).

4. Faktor Fisik dan Lingkungan

Ada dua konsep yang harus dipahami dalam faktor fisik dan lingkungan, yaitu site dan situasi (*situation*). Pengertian tentang site adalah semua sifat atau karakter internal dari

suatu persil atau daerah tertentu, termasuk di dalamnya adalah ukuran (size), bentuk, topografi dan semua keadaan fisik pada persil tanah. Sedangkan yang dimaksud dengan situasi (situation) ialah yang berkenaan dengan sifat-sifat eksternalnya. Situasi suatu tempat berkaitan erat dengan relasi tempat itu dengan tempat-tempat di sekitarnya pada suatu ruang geografi yang sama. Termasuk dalam pengertian situasi adalah aksesibilitas (jarak ke pusat pertokoan (CBD), jarak ke sekolah jarak ke rumah sakit, dan lain-lain), tersedianya sarana dan prasarana (utilitas kota) seperti jaringan transportasi, sambungan telepon, listrik, air minum dan sebagainya. Site mempengaruhi nilai tanah karena “sumberdaya”-nya, sedangkan situasi mempengaruhi nilai tanah karena kemudahan atau kedekatannya (aksesibilitas) dengan “sumberdaya” yang lain di sekitarnya.

Menurut Abd. Rahman M. Noor, seperti yang dikutip oleh [8] dalam jurnal Ekonomi Pembangunan, Penilaian adalah suatu penaksiran dan pendapat atas nilai dari suatu harta tanah/kekayaan oleh seorang penilai yang didasari interpretasi dari faktor-faktor dan keyakinan pada waktu atau tanggal tertentu. Sedangkan menurut Wolcott seperti yang dikutip oleh [8], mengemukakan bahwa konsep nilai ditimbulkan karena adanya faktor-faktor ekonomi sebagai berikut:

- a. Kegunaan (*utility*), yaitu kemampuan suatu benda untuk memuaskan keinginan, kebutuhan dan selera manusia, misalnya tanah yang dapat dibangun rumah di atasnya sebagai tempat tinggal manusia. Kegunaan suatu properti tergantung pada karakteristiknya, seperti ukuran (luas tanah atau bangunan), desain bangunan, aksesibilitas, lokasi, hak kepemilikan dan bentuk lain dari kegunaan yang berpengaruh pada nilai properti.
- b. Kelangkaan (*scarcity*), yaitu suatu barang yang tersedia dalam jumlah yang terbatas akan menjadikan benda tersebut bernilai atau dapat juga dikatakan ketersediaan atau penawaran suatu komoditas relatif terhadap permintaannya.
- c. Keinginan (*desire/demand*), bahwa permintaan terhadap suatu benda menunjukkan benda tersebut bernilai atau harapan pembeli terhadap suatu komoditas untuk dapat memuaskan kebutuhan hidupnya atau keinginan individunya.
- d. Daya beli efektif (*effective purchasing power*), adalah kemampuan seseorang secara individu atau kelompok untuk berpartisipasi di pasar dalam memperoleh suatu komoditi, ditukar dengan sejumlah uang tertentu atau barang lain yang setara dengannya.

Menurut Eldred, seperti yang dikutip oleh [8] dalam jurnal Ekonomi Pembangunan, faktor-faktor yang menentukan nilai ekonomi dari suatu properti tanah adalah:

- a. Permintaan yang menunjukkan keinginan dan kemampuan seseorang untuk membeli atau menyewa suatu properti.
- b. Kegunaan yang menunjukkan manfaat dari properti subyek yang dapat memberikan kepuasan pada konsumen.
- c. Kelangkaan yang menunjukkan kuantitas dan kualitas dari properti lain yang bersaing dengan properti subyek yang bersangkutan.
- d. *Transferability* yaitu, menunjukkan proses pengalihan hak-hak properti dari satu pihak ke pihak lain melalui jual beli, sewa dan kontrak.

2.3 Algoritma *K-Nearest Neighbour*

Algoritma *K-Nearest Neighbor* (KNN) adalah sebuah metode untuk melakukan klasifikasi terhadap objek berdasarkan data pembelajaran yang jaraknya paling dekat dengan objek tersebut. Data pembelajaran diproyeksikan ke ruang berdimensi banyak, dimana masing-masing dimensi merepresentasikan fitur dari data. Ruang ini dibagi menjadi bagian-bagian

berdasarkan klasifikasi data pembelajaran. Sebuah titik pada ruang ini ditandai kelas c jika kelas c merupakan klasifikasi yang paling banyak ditemui pada k buah tetangga terdekat titik tersebut. Dekat atau jauhnya tetangga biasanya dihitung berdasarkan jarak Euclidean dengan rumus seperti pada persamaan (1),

$$distance = \sqrt{\sum_{i=1}^n (X_{training}^i - X_{testing})^2} \quad , \quad (1)$$

dengan

- $X_{training}^i$: data training ke- i ,
- $X_{testing}$: data testing,
- i : record (baris) ke- i dari tabel,
- n : jumlah data training.

Pada fase pembelajaran, algoritma ini hanya melakukan penyimpanan vektor-vektor fitur dan klasifikasi dari data pembelajaran. Pada fase klasifikasi, fitur-fitur yang sama dihitung untuk data test (yang klasifikasinya tidak diketahui). Jarak dari vektor yang baru ini terhadap seluruh vektor data pembelajaran dihitung, dan sejumlah K buah yang paling dekat diambil. Titik yang baru klasifikasinya diprediksikan termasuk pada klasifikasi terbanyak dari titik-titik tersebut. Nilai K yang terbaik untuk algoritma ini tergantung pada data. Secara umum, nilai K yang tinggi akan mengurangi efek *noise* pada klasifikasi, tetapi membuat batasan antara setiap klasifikasi menjadi lebih kabur. Nilai K yang bagus dapat dipilih dengan optimasi parameter, misalnya dengan menggunakan *cross-validation*. Kasus khusus di mana klasifikasi diprediksikan berdasarkan data pembelajaran yang paling dekat (dengan kata lain, $K = 1$) disebut algoritma *nearest neighbour*.

Ketepatan algoritma KNN sangat dipengaruhi oleh ada atau tidaknya fitur-fitur yang tidak relevan, atau jika bobot fitur tersebut tidak setara dengan relevansinya terhadap klasifikasi.

Algoritma KNN memiliki beberapa kelebihan yaitu ketangguhan terhadap *training data* yang memiliki banyak *noise* dan efektif apabila *training data*-nya besar. Sedangkan, kelemahan KNN adalah KNN perlu menentukan nilai dari parameter K (jumlah dari tetangga terdekat), *training* berdasarkan jarak tidak jelas mengenai jenis jarak apa yang harus digunakan dan atribut mana yang harus digunakan untuk mendapatkan hasil terbaik, dan biaya komputasi cukup tinggi karena diperlukan perhitungan jarak dari tiap *query instance* pada keseluruhan *training sample*.

K-Nearest Neighbour (KNN) adalah suatu metode yang menggunakan algoritma *supervised* dimana hasil dari *query instance* yang baru diklasifikasikan berdasarkan mayoritas dari kategori pada KNN. Tujuan dari algoritma ini adalah mengklasifikasikan obyek baru berdasarkan atribut dan *training sample*. *Classifier* tidak menggunakan model apapun untuk dicocokkan dan hanya berdasarkan pada memori. Diberikan titik *query*, akan ditemukan sejumlah K obyek atau (titik *training*) yang paling dekat dengan titik *query*. Klasifikasi menggunakan *voting* terbanyak diantara klasifikasi dari K obyek. Algoritma KNN menggunakan klasifikasi ketetanggaan sebagai nilai prediksi dari *query instance* yang baru.

Sebagai ilustrasi dari penerapan algoritma KNN adalah misalnya terdapat data hasil survei dengan kuesioner, untuk meminta pendapat orang pada pengujian dengan dua atribut (ketahanan asam dan kekuatan), untuk mengklasifikasikan apakah suatu kertas tisu berkualitas baik atau tidak [9]. Berikut ini adalah empat sampel pelatihan (data *training*).

Tabel 1. Tabel Data *Training*.

$X_1 = \text{Asam Durabilitas}$ (detik)	$X_2 = \text{Kekuatan}$ (Kg / meter persegi)	Klasifikasi
7	7	Buruk
7	4	Buruk
3	4	Baik
1	4	Baik

Sebagai kasus, misalnya saat ini pabrik kertas telah menghasilkan jaringan baru yang lulus uji laboratorium dengan $X_1=3$ dan $X_2=7$. Untuk menebak klasifikasi jaringan baru ini maka dilakukan perhitungan dengan menggunakan algoritma KNN.

Adapun langkah-langkah untuk menghitung K tetangga terdekat dengan algoritma KNN adalah sebagai berikut.

- Tentukan parameter K (jumlah tetangga terdekat). Misalkan $K = 3$
- Hitung jarak antara permintaan (data testing) dan contoh-contoh latihan semua (data training). Data training yang akan dihitung kedekatannya mempunyai koordinat (3,7).

Tabel 2. Tabel Perhitungan Jarak.

$X_1 = \text{Asam Durabilitas}$ (detik)	$X_2 = \text{Kekuatan}$ (Kg / meter persegi)	Square Jarak ke contoh permintaan (3, 7)
7	7	$(7-3)^2 + (7-7)^2 = 16$
7	4	$(7-3)^2 + (4-7)^2 = 25$
3	4	$(3-3)^2 + (4-7)^2 = 9$
1	4	$(1-3)^2 + (4-7)^2 = 13$

- Urutkan jarak dan menentukan tetangga terdekat berdasarkan jarak terdekat ke- K .

Tabel 3. Tabel Urutan Peringkat Jarak Minimum.

$X_1 = \text{Asam Durabilitas}$ (detik)	$X_2 = \text{Kekuatan}$ (Kg / meter persegi)	Square Jarak ke contoh permintaan (3, 7)	Peringkat jarak minimum	Apakah termasuk dalam- tetangga terdekat 3?
7	7	$(7-3)^2 + (7-7)^2 = 16$	3	Ya
7	4	$(7-3)^2 + (4-7)^2 = 25$	4	Tidak
3	4	$(3-3)^2 + (4-7)^2 = 9$	1	Ya
1	4	$(1-3)^2 + (4-7)^2 = 13$	2	Ya

- Kumpulkan kategori Y dari baris tetangga terdekat. Pada baris kedua kategori tetangga terdekat (Y) tidak dimasukkan karena data tersebut peringkatnya lebih dari 3 tetangga terdekat.

Tabel 4. Kumpulan Kategori Y Tetangga Terdekat.

$X_1 = \text{Asam Durabilitas}$ (detik)	$X_2 = \text{Kekuatan}$ (Kg / meter	Square Jarak ke contoh permintaan	Peringkat jarak minimum	Apakah termasuk dalam	$Y =$ Kategori terdekat

	persegi)	(3, 7)		tetangga 3- Terdekat?	Tetangga
7	7	$(7-3)^2 + (7-7)^2 = 16$	3	Ya	Buruk
7	4	$(7-3)^2 + (4-7)^2 = 25$	4	Tidak	-
3	4	$(3-3)^2 + (4-7)^2 = 9$	1	Ya	Baik
1	4	$(1-3)^2 + (4-7)^2 = 13$	2	Ya	Baik

- e. Gunakan mayoritas sederhana dari kategori tetangga terdekat sebagai nilai prediksi contoh *query*.

Dari tabel di atas diperoleh dua kertas tisu baru berkualitas baik dan satu kertas tisu baru berkualitas buruk. Karena tetangga terdekat yang didapat lebih banyak yang berkualitas baik, maka dapat disimpulkan bahwa kertas tisu baru yang lulus uji laboratorium dengan $X_1=3$ dan $X_2=7$ adalah termasuk dalam kategori **baik**.

3. Metodologi

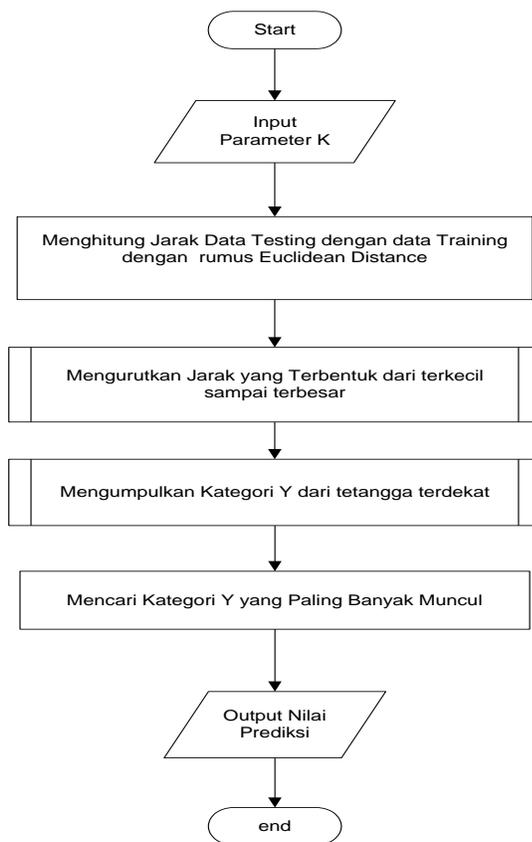
Metode yang digunakan adalah menggunakan metode rekayasa dengan pendekatan model waterfall yang meliputi tahap analisis, desain, implementasi dan uji coba yang digabungkan dengan metode data mining yang dikenal dengan istilah *Cross-Industry Standard Process for Data Mining* (CRISP-DM). Pada tahap analisis dilakukan survei pendahuluan mengenai data-data yang diduga berpengaruh terhadap harga jual tanah. Data yang dikumpulkan berupa data sekunder yaitu berupa lebar jalan, jarak pertokoan, kondisi jalan, kepadatan penduduk, bebas banjir, transportasi umum, jaringan telepon, saluran PDAM, jarak sekolah, jarak rumah sakit, saluran listrik, surat hak milik, ukuran tanah, supply demand, ada bangunan, arah mata angin, dan nilai tanah (harga jual tanah). Variabel-variabel diatas selanjutnya dijadikan acuan untuk pengumpulan data.

Setelah melakukan analisis sistem, selanjutnya dilakukan perancangan proses, input, output, database, *graphical user interface* (GUI), dan program yang akan dibuat untuk melakukan prediksi. Secara umum, desain proses dari aplikasi yang dibuat seperti dalam Gambar 1.



Gambar 1. *Context Diagram*.

Prosedur aplikasi untuk diagram di atas diberikan pada Gambar 2 berikut.



Gambar 2. Prosedur Aplikasi dengan KNN.

4. Hasil dan Pembahasan

Pada bagian ini akan dibahas mengenai hasil implementasi dan pembahasan hasil uji coba program aplikasi. Aplikasi dibuat dengan struktur menu seperti pada Gambar 3.



Gambar

3. Struktur Menu Aplikasi.

Proses pada data pelatihan (training) dan data pengujian (testing) merupakan inti pembahasan dari aplikasi yang dibuat. Data training digunakan untuk menyimpan kasus lama yang sudah diketahui nilai tanahnya, sehingga ketika ada data baru yang dimasukkan (disebut dengan data testing) dapat diprediksi nilai tanahnya berdasarkan ukuran kedekatan. Sebagai contoh data training yang

Wiyli Yustanti

digunakan seperti tampak pada Gambar 4. Di dalam data training tersebut terlihat nilai dari masing-masing atribut yang berkaitan dengan sebuah data tanah pada lokasi tertentu. Dalam penelitian ini mengambil sampel data tanah di wilayah Surabaya sebanyak 100 sampel.

ID Training	Lebar Jalan	Jarak Toko	Kondisi Jalan	Penduduk	Bebas Banjir	Transportasi Umum	Telepon
1	> 9 meter	3 < X2 < 6 km	Cukup Baik	Sedang	Bebas Banjir	Tidak Ada Transportasi	Tidak Ada Jaringan
2	3 <= X1 < 6 meter	<= 3 km	Cukup Baik	Jarang Penduduk	Banjir	Ada Transportasi	Tidak Ada Jaringan
3	< 3 meter	>= 6 km	Cukup Baik	Padat	Banjir	Tidak Ada Transportasi	Ada Jaringan
4	3 <= X1 < 6 meter	<= 3 km	Baik	Jarang Penduduk	Banjir	Ada Transportasi	Ada Jaringan
5	6 <= X1 < 9 meter	3 < X2 < 6 km	Cukup Baik	Jarang Penduduk	Bebas Banjir	Ada Transportasi	Tidak Ada Jaringan
6	< 3 meter	3 < X2 < 6 km	Cukup Baik	Sedang	Bebas Banjir	Ada Transportasi	Ada Jaringan
7	> 9 meter	3 < X2 < 6 km	Kurang Baik	Sedang	Bebas Banjir	Tidak Ada Transportasi	Ada Jaringan
8	3 <= X1 < 6 meter	>= 6 km	Cukup Baik	Sedang	Banjir	Ada Transportasi	Ada Jaringan
9	3 <= X1 < 6 meter	3 < X2 < 6 km	Baik	Padat	Banjir	Ada Transportasi	Ada Jaringan
10	6 <= X1 < 9 meter	<= 3 km	Baik	Padat	Bebas Banjir	Ada Transportasi	Ada Jaringan
11	> 9 meter	>= 6 km	Baik	Sedang	Bebas Banjir	Tidak Ada Transportasi	Ada Jaringan
12	3 <= X1 < 6 meter	>= 6 km	Cukup Baik	Sedang	Banjir	Ada Transportasi	Ada Jaringan
13	< 3 meter	<= 3 km	Kurang Baik	Jarang Penduduk	Bebas Banjir	Ada Transportasi	Ada Jaringan
14	3 <= X1 < 6 meter	>= 6 km	Cukup Baik	Sedang	Bebas Banjir	Ada Transportasi	Ada Jaringan
15	6 <= X1 < 9 meter	>= 6 km	Baik	Sedang	Banjir	Ada Transportasi	Ada Jaringan
16	> 9 meter	3 < X2 < 6 km	Cukup Baik	Padat	Banjir	Tidak Ada Transportasi	Ada Jaringan
17	< 3 meter	3 < X2 < 6 km	Baik	Padat	Banjir	Ada Transportasi	Ada Jaringan
18	6 <= X1 < 9 meter	>= 6 km	Baik	Sedang	Banjir	Ada Transportasi	Ada Jaringan
19	> 9 meter	<= 3 km	Cukup Baik	Sedang	Bebas Banjir	Ada Transportasi	Ada Jaringan
20	3 <= X1 < 6 meter	3 < X2 < 6 km	Cukup Baik	Padat	Bebas Banjir	Ada Transportasi	Ada Jaringan
23	< 3 meter	3 < X2 < 6 km	Cukup Baik	Sedang	Bebas Banjir	Tidak Ada Transportasi	Tidak Ada Jaringan

Gambar 4. Sampel Data *Training*.

Untuk melakukan pengujian data *testing* maka dilakukan input data *testing* dan memasukkan jumlah data terdekat (K) yang dipertimbangkan. Output dari hasil perhitungan jarak data membandingkan nilai atribut pada data *testing* dengan nilai atribut pada semua data *training* akan ditampilkan pada dialog disebelah kiri.

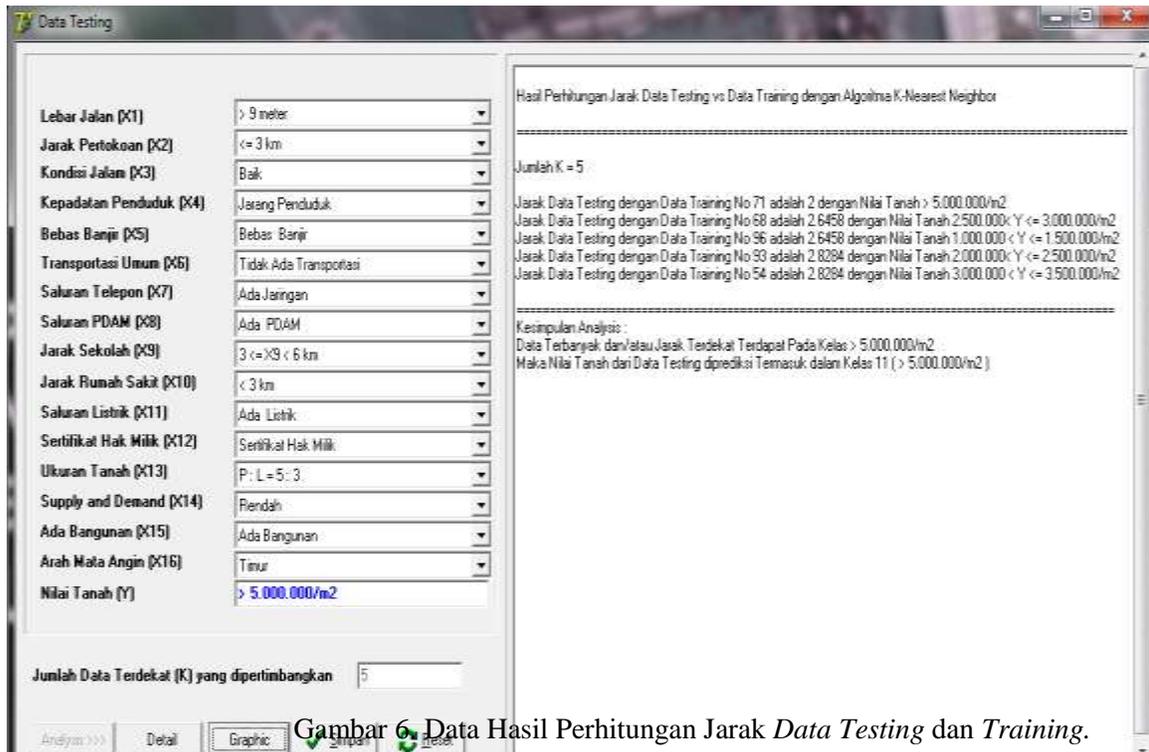
NO	X1	X2	X3	X4	X5	X6	X7	X8	X9	X10	X11	X12	X13
1	4	2	2	2	1	1	2	2	1	2	1	2	

NO	X1	X2	X3	X4	X5	X6	X7	X8	X9	X10	X11	X12	X13
1	4	2	2	2	1	1	1	2	1	2	1	2	
2	2	1	2	1	2	2	1	1	2	2	1	1	
3	1	3	2	3	2	1	2	2	3	2	2	1	
4	2	1	3	1	2	2	2	2	2	3	2	2	
5	3	2	2	1	1	2	1	2	3	3	2	2	
6	1	2	2	2	1	2	2	2	3	2	2	3	
7	4	2	1	2	1	1	2	1	3	2	2	3	
8	2	3	2	2	2	2	2	1	2	3	2	1	
9	2	2	3	3	2	2	2	1	2	1	2	4	

dC0	dC7	dC8	dC9	dC10	dC11	dC12	dC13	dC14	dC15	dC16	Distance
0	-1	0	0	0	0	0	0	0	0	0	1
1	-1	-1	1	0	0	-1	1	0	1	-2	4.2426
0	0	0	2	0	1	-1	1	1	1	0	4.5826
1	0	0	1	1	1	0	2	1	0	0	4.1231
1	-1	0	2	1	1	0	2	0	1	-1	4
1	0	0	2	0	1	1	2	0	1	0	4.5826
0	0	-1	2	0	1	1	3	1	1	-1	4.4721
1	0	-1	1	1	1	3	1	1	0	-1	4.2426
1	0	-1	1	1	1	3	1	1	0	-1	4.5826

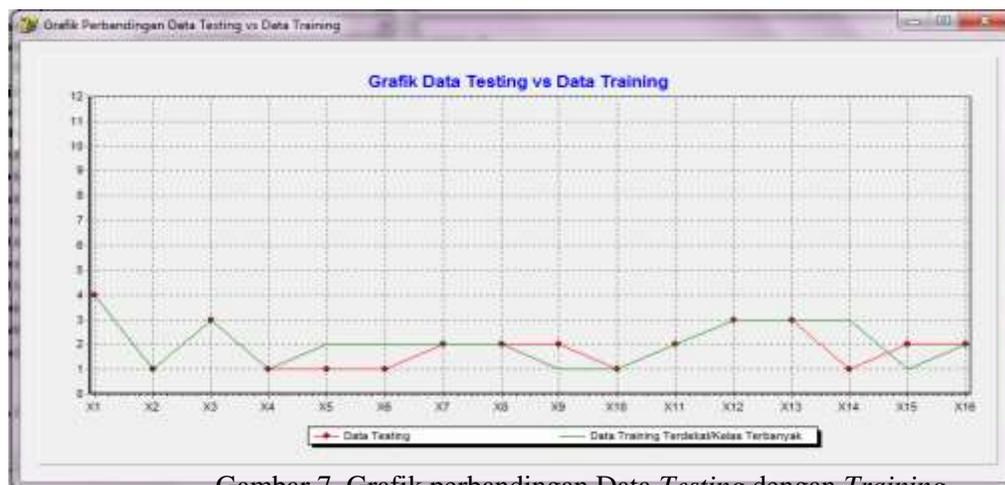
Gambar 5. Proses Analisa Data *Testing*.

Untuk melihat lebih detail perhitungan data *training* dan data *testing* serta jaraknya, dapat dilihat pada menu *Detail*. Selanjutnya akan muncul form yang menampilkan semua data yang dibutuhkan dalam perhitungan jarak kedekatan antara data baru dan data lama. Contoh hasil perhitungan jarak dari data *testing* dan data *training* dapat dilihat pada Gambar 6.



Gambar 6. Data Hasil Perhitungan Jarak *Data Testing* dan *Training*.

Sedangkan untuk melihat grafik perbandingan data *testing* dengan *data training* diberikan fasilitas untuk melihat grafiknya. Sebagai contoh perbandingan tersebut dapat dilihat pada Gambar 7.



Gambar 7. Grafik perbandingan *Data Testing* dengan *Training*.

Setelah aplikasi siap digunakan maka dilakukan evaluasi untuk mengukur tingkat presisi dari hasil prediksi nilai tanah di wilayah Surabaya. Evaluasi tersebut dapat dilihat pada tabel dibawah ini.

Tabel 5. Tabel Tingkat Presisi Prediksi Nilai Tanah di Wilayah Surabaya.

No	Wilayah/Alamat	Harga Sebenarnya	Harga Prediksi	Keputusan
1.	Gayungan, A. Yani	10.000.000/m ²	> 5.000.000/m ²	Benar
2.	Nginden, Semolo	4.350.000/m ²	> 5.000.000/m ²	Salah
3.	Jl. Kedungdoro	6.200.000/m ²	> 5.000.000/m ²	Benar
4.	Prapen / Jemursari	6.500.000/m ²	> 5.000.000/m ²	Benar
5.	Jl. Embong Wungu	8.000.000/m ²	> 5.000.000/m ²	Benar
6.	Jurang Kuping,	530.000/m ²	500.000<Y=<1jt/m ²	Benar
7.	Jl. Pakis	1.960.000/m ²	1,5 jt<Y<=2 jt/m ²	Benar
8.	Keputih, Sukolilo	1.500.000/m ²	1,5 jt<Y<=2 jt/m ²	Salah
9.	Tenggilis, Mejoyo	1.600.000/m ²	1,5 jt<Y<=2 jt/m ²	Benar
10.	Wonorejo Asri	485.000/m ²	400.000<Y<=500.000 /m ²	Benar

Berdasarkan hasil prediksi data di atas maka tingkat akurasi dari prediksi ini dapat dikatakan sebesar 80%. Tingkat akurasi ini dapat diperbaiki dengan menambah jumlah data training, sehingga kasus baru yang akan diprediksi harganya sudah ada nilai acuannya di dalam database.

5. Penutup

Berdasarkan penelitian ini diperoleh bahwa aplikasi yang dihasilkan memiliki tingkat pemrosesan yang cukup lama karena data *testing* (data baru) dibandingkan satu persatu dengan data *training* (data lama). Hal ini karena jumlah atribut yang digunakan untuk memprediksi nilai tanah cukup banyak, yaitu 16 variabel yang meliputi lebar jalan, kondisi jalan, kondisi banjir, jarak antara tanah dengan pertokoan, sekolah dan rumah sakit, terdapat saluran telepon, PDAM, listrik dan masih banyak lagi yang lain. Tingkat presisi dari prediksi yang dihasilkan sekitar 80%. Untuk meningkatkan akurasinya, maka diperlukan tambahan data sampel untuk proses pelatihan yang lebih banyak. Selain itu juga dapat dilakukan penelitian lanjutan untuk melakukan prediksi dengan algoritma yang lainnya seperti analisis regresi, *decscion tree* dan sebagainya.

Daftar Pustaka

- [1] Alexander M.H., 2004. *Pemrograman Database Menggunakan Delphi 7.0 dengan Metode ADO*. PT. Elex Media Komputindo, Jakarta.
- [2] Kristanto Y., 2010. *Daftar Perkiraan Harga-harga Tanah Pasar Sekunder di Surabaya Per 01 Maret 2010*, (Online). Sumber: <http://www.PropertiGuide.com> [diakses 01 Maret 2010]
- [3] Kusrini dan Luthfi E.T., 2009. *Algoritma Data Mining*. CV. Andi Offset, Yogyakarta.

- [4] Northam R.M., 1975. *Urban Geography*. John Wiley & Sons, Inc., New York.
- [5] Pearce D.W., dan Turner K.R., 1990. *Economics Of Natural Resources and The Enviroment*. The John Hopkins University, Baltimore.
- [6] Saputra H., 2008. Faktor-faktor yang mempengaruhi pembangunan rumah di Daerah Istimewa Yogyakarta (1991-2005). *Skripsi*, tidak dipublikasikan. Universitas Islam Indonesia, Yogyakarta.
- [7] Soekanto, 1994. *Pengantar Sosiologi Edisi VI*. Radjawali Press.
- [8] Sutawijaya A., 2004. Analisis faktor–faktor yang mempengaruhi nilai tanah sebagai dasar penilaian Nilai Jual Obyek Pajak (NJOP) PBB di Kota Semarang. *Jurnal Ekonomi Pembangunan*, Vol. 9 No. 1, Juni 2004, hal. 65-78.
- [9] Teknomo K., 2010. *K-Nearest Neighbours Tutorial, K-Tetangga Terdekat Tutorial*, (Online). Sumber: <http://people.revoledu.com/kardi/tutorial/KNN/> [diakses 4 Maret 2010]
- [10] Wolcott R.C., 1987. The appraisal of Real Estate American Institute of Real Estate Appraiser. *North Michigan*, Chicago Illinois, hal. 22-63.