

PROPOSAL TUGAS AKHIR

Penelitian Komparatif : Prediksi Harga Rumah Menggunakan Teknik *Linear Regression* dan *Ridge Regression*

Disusun oleh:

1. 191111053 Tubagus Mochammad Eza Rizki Firdaus
2. 191111054 Muhammad Husni Abdillah
3. 191111044 Ardhiansyah Kurniawan
4. 191111052 Hikmah



PROGRAM STUDI TEKNIK INFORMATIKA

SEKOLAH TINGGI INFORMATIKA & KOMPUTER INDONESIA

STIKI MALANG

2021

LEMBAR PERSETUJUAN

PROPOSAL TUGAS AKHIR

JUDUL

Penelitian Komparatif : Prediksi Harga Rumah Menggunakan Teknik *Linear Regression* dan *Ridge Regression*

Penyusun:

1. 191111053 Tubagus Mochammad Eza Rizki Firdaus
2. 191111054 Muhammad Husni Abdillah
3. 191111044 Ardhiansyah Kurniawan
4. 191111052 Hikmah

Periode Pengajuan Gasal 2021

Program Studi Teknik Informatika

Malang, 11 Januari 2022

Mengetahui,

Ketua Program Studi TI,

Pembimbing,

Yekti Asmoro K., S.Si. M.A.B

Rakhmad Maulidi, M. Kom

Daftar Isi

LEMBAR PERSETUJUAN	2
Daftar Isi	3
Daftar Tabel	4
Daftar Gambar	5
Bab I Pendahuluan	6
1.1 Latar Belakang	6
1.2 Rumusan Masalah / hipotesis penelitian	7
1.3 Manfaat	7
1.4 Batasan Masalah	7
1.5 Sistematika Penulisan	7
Bab II Tinjauan Pustaka	8
2.1 Tinjauan Empiris	8
2.2 Tinjauan Teoritis	10
2.3 Hipotesa	10
Bab III Metodologi Penelitian	11
3.1 Jenis Penelitian	11
3.2 Kerangka Konsep Penelitian	11
3.3 Populasi dan Sampel Penelitian	12
3.3.1 Populasi	12
3.3.2 Sampel	12
3.3.3 Karakteristik Responden	12
3.4 Pengumpulan Data	12
3.4.1 Jenis data	12
3.4.2 Model Pengumpulan Data	12
3.4.3 Karakteristik Responden	13
3.5 Definisi Operasional Variabel	13
3.6 Rancangan Pengujian	13
3.7 Teknik Analisis Data	13
Daftar Pustaka	14

Daftar Tabel

Daftar Gambar

<i>Gambar 3.1. Flowchart penelitian.....</i>	12
--	-----------

Bab I Pendahuluan

1.1 Latar Belakang

Rumah adalah salah satu dari kebutuhan *premier* dari masyarakat yang tidak dapat dihindari, karena merupakan tempat untuk berlindung dan melepas penat setelah beraktivitas. Tidak kalah dengan emas dan saham, rumah juga bisa digunakan untuk berinvestasi, dikarenakan perubahan harganya yang dapat berubah sewaktu-waktu dan semakin banyaknya *demand*, terlebih lagi pada daerah yang padat penduduk seperti perkotaan dan dekat dengan perkantoran maupun fasilitas umum. Hal tersebut pasti akan mempengaruhi harga rumah dengan signifikan (Andi Saiful, 2021).

Seiring dengan perkembangan, prediksi digunakan sebagai alat bantu dan bahkan menjadi pertimbangan dalam pengambilan keputusan (Ayuni G. N., 2019). Dari penelitian yang dilakukan oleh A. Saiful (2021) dengan penelitian yang berjudul “Prediksi Harga Rumah Menggunakan *Web Scrapping dan Machine Learning* dengan Algoritma *Linear Regression*” mengatakan bahwa *linear regression* dapat memberikan hasil keakuratan prediksi harga rumah dengan baik. Dengan tingkat akurasi sebesar 88%, namun memiliki nilai galat atau RMSE (*Root Mean Square Error*) cukup tinggi sebesar 259171,91 (Andi Saiful, 2021).

Pada penelitian ini kami mencoba untuk menggunakan beberapa teknik regresi lain yang cocok digunakan untuk permasalahan tersebut. Teknik regresi yang akan kami bandingkan adalah *Linear Regression* dan *Ridge Regression*. *Linear regression* digunakan untuk mengetahui pengaruh antara

satu atau beberapa variabel (variabel bebas) terhadap satu buah variabel (Saputro, 2018). *Ridge Regression* adalah metode yang berdasar dari regresi linear yang modifikasi lagi untuk mengatasi masalah *multikolinearitas* pada data, atau keadaan dimana adanya hubungan saling berpengaruh antar variabel bebas yang menyebabkan data hasil prediksi bernilai jauh dari data yang sebenarnya.

1.2 Rumusan Masalah / hipotesis penelitian

Berdasarkan latar belakang di atas maka perlu dilakukan pengujian tingkat akurasi dengan melihat RMSE yang dihasilkan pada dua model yaitu *Linear Regression* dan *Ridge Regression* dalam memprediksi harga rumah. Adapun rumusan masalah dari penelitian ini adalah “Sejauh mana perbedaan akurasi pada metode linear regression dan ridge regression dalam memprediksi harga rumah”.

1.3 Manfaat

Manfaat dari penelitian ini adalah untuk mempelajari dan mengetahui perbandingan dan algoritma mana yang dinilai lebih akurat berdasarkan pada tingkat RMSE yang dihasilkan.

1.4 Batasan Masalah

1.5 Sistematika Penulisan

Bab II Tinjauan Pustaka

2.1 Tinjauan Empiris

Adapun yang menjadi landasan penelitian terdahulu dalam penelitian ini adalah sebagaimana hasil penelitian yang dilakukan dibawah ini, yakni oleh:

Penelitian yang dilakukan oleh Sefto (2016) dengan penelitian “Prediksi Harga Tanah Menggunakan Algoritma Linear Regression”. Mengatakan bahwa *linear regression* dapat memberikan hasil pengukuran tingkat akurasi yang tinggi dan RMSE (*Root Mean Square Error*) terendah dengan nilai 77072562,116.

Penelitian yang dilakukan oleh A. Saiful (2021) dengan penelitian yang berjudul “Prediksi Harga Rumah Menggunakan *Web Scrapping dan Machine Learning* dengan Algoritma *Linear Regression*” mengatakan bahwa *linear regression* dapat memberikan hasil keakuratan prediksi harga rumah dengan baik. Dengan tingkat akurasi sebesar 77%, namun memiliki nilai galat atau RMSE (*Root Mean Square Error*) cukup tinggi sebesar 259171,91 (Andi Saiful, 2021).

Penelitian yang dilakukan oleh Ohhyver (2011) yaitu “Metode Regresi Ridge untuk Mengatasi Kasus *Multikolinear*” . Menyatakan bahwa *ridge regression* dapat mendekati nilai parameter yang sebenarnya dengan nilai RMSE terkecil. Adapun *ridge regression* adalah metode yang berdasarkan *linear regression* yang dimodifikasi dengan ditambahkannya konstanta bias c

untuk mengatasi korelasi yang tinggi antara beberapa variabel bebas (Ohlyver, 2011).

No	Nama Peneliti	Judul	Tujuan	Manfaat	Hasil
1.	Sefto Pratama (2016)	Prediksi Harga Tanah Menggunakan Algoritma Linear Regression	Mengetahui keakuratan metode <i>Linear Regression</i> dalam memprediksi harga tanah yang mendekati harga ideal	Mempelajari dan memfasilitasi pemahaman tentang metode <i>Linear Regression</i>	Algoritma <i>Linear Regression</i> dapat memberikan hasilakurasi atau pengukuran keakuratan dari prediksi penentuan dengan baik
2.	Andi Saiful (2021)	Prediksi Harga Rumah Menggunakan <i>Web Scrapping</i> dan <i>Machine Learning</i> dengan Algoritma <i>Linear Regression</i>	Membuat model prediksi harga rumah menggunakan metode <i>machine learning</i> dan algoritma <i>linear regression</i>	Mempelajari dan mengetahui tentang penerapan <i>machine learning</i> menggunakan algoritma <i>linear regression</i> yang digunakan	Hasil prediksi harga rumah menggunakan <i>linear regression</i> memberikan hasil keakuratan yang baik. Dengan nilai akurasi 88% namun dengan galat

				untuk memprediksi harga rumah	yang cukup tinggi sebesar 259171,91
3.	Margaretha Ohyver (2011)	Metode Regresi Ridge untuk Mengatasi Kasus Multikolinear	Memperoleh persamaan regresi Ridge yang dapat mengatasi kasus multikolinear	Memberikan pengetahuan mengenai cara mengatasi adanya multikolinear serta penerapan regresi ridge pada data.	

2.2 Tinjauan Teoritis

2.3 Hipotesa

Dari uraian tersebut dapat di tarik penduga bahwa *ridge regression* mempunyai tingkat keakuratan yang lebih tinggi daripada *linear regression*. Untuk itu hipotesa dari penelitian ini adalah “Apakah *ridge regression* lebih akurat daripada *linear regression* dalam memprediksi harga rumah?”

Hipotesis dalam penelitian ini adalah sebagai berikut:

H₀ : *Ridge regression* lebih akurat daripada *linear regression* dalam memprediksi harga rumah.

H_a : *Ridge regression* tidak lebih akurat daripada *linear regression* dalam memprediksi harga rumah.

Bab III Metodologi Penelitian

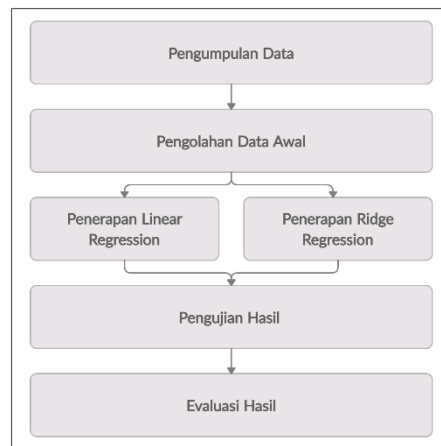
3.1 Jenis Penelitian

Penelitian komparatif merupakan sejenis penelitian deskriptif berusaha menemukan jawaban sebab-akibat yang mendasar, dengan menganalisis faktor-faktor yang menyebabkan munculnya atau berkembangnya suatu fenomena tertentu (Nazir, 2005).

Penelitian ini termasuk dalam kategori penelitian deskriptif. Objek dari penelitian ini berfokus pada tingkat keakuratan dilihat dari nilai RMSE yang dihasilkan. Sedangkan subjek penelitian ini adalah nilai prediksi harga rumah menggunakan metode *linear regression* dan *ridge regression*.

3.2 Kerangka Konsep Penelitian

Penelitian ini dilakukan dengan berbagai langkah kerja yang sistematis sehingga mendapatkan hasil yang optimal. Langkah kerja penelitian merupakan serangkaian prosedur dan langkah-langkah dalam melakukan penelitian yang terstruktur secara sistematis dan terarah agar tujuan dari penelitian bisa tercapai dengan baik. Adapun langkah-langkah yang digunakan dalam penelitian ini dilihat dalam *flowchart* penelitian.



Gambar 3.1. Flowchart penelitian

3.3 Populasi dan Sampel Penelitian

3.3.1 Populasi

3.3.2 Sampel

3.3.3 Karakteristik Responden

3.4 Pengumpulan Data

3.4.1 Jenis data

Data yang digunakan dalam penelitian ini adalah data penjualan rumah di Amerika Serikat yang disediakan oleh platform *Kaggle*. Data set ini mempunyai 80 variabel yang dianggap berpengaruh terhadap harga rumah. Untuk penelitian ini hanya digunakan 6 variabel yaitu rata-rata pendapatan penduduk pada suatu daerah, rata-rata usia rumah pada suatu daerah, rata-rata jumlah kamar setiap rumah pada suatu daerah, populasi penduduk pada suatu daerah, alamat rumah, dan harga rumah.

3.4.2 Model Pengumpulan Data

Dengan data yang disediakan oleh *Kaggle* dalam suatu kompetisi yang berjudul “

3.4.3 Karakteristik Responden

3.5 Definisi Operasional Variabel

3.6 Rancangan Pengujian

Dalam penelitian ini menggunakan Project Jupyter yang merupakan aplikasi open-source digunakan untuk membuat dokumen yang dapat menjalankan code, visualisasi, menyelesaikan persamaan, dan narasi berupa teks. Terdapat *tools* untuk *data cleaning*, tranformasi data, simulasi nilai numerik, permodelan menggunakan statistik, visualisasi data dan *tools* untuk *machine learning*.

3.7 Teknik Analisis Data

Daftar Pustaka

- Andi Saiful, S. A. (2021). Prediksi Harga Rumah Menggunakan Web Scrapping dan Machine Learning dengan Algoritma Linear Regression. *Jurnal Teknik Informatika dan Sistem Informasi*, 8, 41-50.
- Ayuni G. N., F. D. (2019). Penerapan Metode Regresi Linear untuk Prediksi Penjualan Properti pada PT XYZ. *Jurnal Telematika*, 14, 79-86.
- Nazir, M. (2005). *Metodologi Penelitian*. Bogor: Ghalia Indonesia.
- Ohyver, M. (2011). Metode Regresi Ridge untuk Mengatasi Kasus Multikolinear. *ComTech: Computer, Mathematics and Engineering Applications*, 451-457.
- Pratama, S. (2016). Prediksi Harga Tanah menggunakan Algoritma Linear Regression. *Technologia*, 7(2), 112-130.
- Saputro, E. W. (2018). *Linear Regression*. medium.com.