

# BAB I

## PENDAHULUAN

### 1.1. Latar Belakang

Pada era modern ini, *real estate* menjadi salah satu industri yang berkembang secara signifikan. Persaingan berkembang seiring dengan bertumbuhnya teknologi, mengharuskan para *developer real estate* untuk mencari cara baru dalam bersaing. Salah satu strategi pemasaran yang efektif adalah dengan melakukan optimasi pada mesin pencari (*Search Engine Optimization/SEO*) untuk meningkatkan visibilitas *website* perusahaan di mesin pencari seperti Google [1].

Pemasaran melalui internet dan optimasi mesin pencari (*SEO*) menjadi opsi penting untuk mempertahankan dan meningkatkan pangsa pasar perusahaan *real estate*. Perusahaan *real estate* yang memiliki *website* yang dioptimalkan untuk mesin pencari, memiliki kesempatan lebih besar untuk muncul web pada halaman pertama hasil pencarian dan meningkatkan visibilitasnya di internet [2].

Teknik *SEO* bekerja dengan meningkatkan visibilitas *website* perusahaan *real estate* di internet dan mendapatkan lebih banyak calon pembeli atau penyewa properti yang potensial. Perusahaan *real estate* dapat melakukan penelitian tentang kata kunci yang paling sering dicari oleh calon pembeli atau penyewa properti untuk dijadikan fokus dalam *website* mereka. Dengan memilih kata kunci yang tepat, *website* perusahaan akan lebih mudah ditemukan oleh calon pembeli atau penyewa properti melalui mesin pencari seperti Google. Selain memilih kata kunci yang tepat, teknik *SEO* lainnya adalah memastikan *website* perusahaan memiliki struktur yang baik dan mudah diakses oleh pengguna. Hal ini dapat dilakukan dengan memperbaiki struktur tautan (*link structure*) dan mengoptimalkan deskripsi meta (*meta description*) pada setiap halaman *website*. Dengan struktur yang baik dan mudah diakses, *website* perusahaan akan lebih mudah ditemukan oleh mesin pencari dan calon pembeli atau penyewa properti [3].

Analisis kata kunci dilakukan dengan mengidentifikasi kata kunci yang relevan dengan produk atau layanan yang ditawarkan oleh perusahaan, serta kata kunci yang paling sering dicari oleh calon pembeli atau penyewa. Kemudian, kata kunci tersebut diurutkan berdasarkan volume pencarian dan tingkat persaingan yang ada di pasar [4]. Adapun dengan mengetahui kata kunci yang memiliki *trend* naik (baik), perusahaan dapat memanfaatkan *trend* tersebut untuk bersaing dengan meningkatkan visibilitas *website*, hal ini mendorong adanya peramalan (*forecasting*) untuk *trend keyword SEO* agar perusahaan dapat bersaing secara efektif.

*Forecasting* atau peramalan adalah suatu proses untuk memprediksi nilai-nilai masa depan suatu variabel berdasarkan data historis yang telah dikumpulkan. Banyak metode peramalan untuk *data time series* yang dapat digunakan seperti *prophet*, *moving average*, *LSTM (long short-term memory)*, *proyeksi holt-winter* dan lain sebagainya. Metode regresi linear dan *ARIMA (autoregressive integrated moving average)* merupakan salah satu metode peramalan yang paling umum digunakan dalam analisis data *time series* [6].

Penelitian ini menggunakan metode regresi linear sederhana dan *ARIMA* karena memiliki beberapa kelebihan seperti pada regresi linear yang sederhana mudah dipahami dan diterapkan. Selain itu, regresi linear juga dapat diinterpretasikan dengan mudah. Koefisien regresi memberikan informasi tentang seberapa besar perubahan dalam variabel dependen yang diakibatkan oleh perubahan satu unit dalam variabel independent. Kelebihan yang paling mempengaruhi pertimbangan pemilihan regresi linear sebagai metode peramalan dalam penelitian ini adalah karena regresi linear sederhana cenderung lebih cepat dalam perhitungan dibandingkan dengan beberapa metode peramalan yang lebih kompleks seperti beberapa model jaringan saraf tiruan.

Metode *ARIMA* juga memiliki banyak kelebihan yang membuatnya dipilih untuk meramalkan *trend keyword SEO* dalam penelitian ini. Salah satu kelebihan dari metode *ARIMA* adalah dapat menangani data *time series* yang memiliki tren (kenaikan atau penurunan secara berkelanjutan dalam jangka

waktu tertentu) dan komponen musiman (fluktuasi yang berulang pada periode waktu tertentu) namun, hal ini akan merubah metode *ARIMA* menjadi *SARIMA* (*seasonal-ARIMA*) dengan memasukan unsur musiman (*seasonal*). Selain itu *ARIMA* juga mampu menangani data yang tidak stasioner dengan menggunakan operasi diferensiasi. Diferensiasi digunakan untuk membuat data stasioner, yaitu data yang memiliki mean dan varians konstan dari waktu ke waktu. Kelebihan yang paling mempengaruhi pertimbangan pemilihan *ARIMA* sebagai metode peramalan dalam penelitian ini adalah karena dapat digunakan sebagai basis untuk pengembangan model peramalan yang lebih kompleks. Dalam beberapa kasus, model *ARIMA* dapat digunakan sebagai awal untuk memahami pola-pola dasar dalam data sebelum menjelajahi model-model peramalan yang lebih canggih.

Dalam penelitian *forecasting* dengan metode regresi linear dan *ARIMA*, langkah awal yang dilakukan adalah mengidentifikasi pola dan tren dari data historis yang telah dikumpulkan. Kemudian dilakukan analisa untuk mendapatkan komponen hyper parameter model *ARIMA*, hyper parameter model *ARIMA* terdiri komponen  $p$ ,  $d$ , dan  $q$  dimana setiap komponen ini merepresentasikan komponen *ARIMA* yaitu *autoregressive*, *integrated*, dan *moving average*. Analisa hyper parameter dilakukan dengan menggunakan analisis statistik, seperti grafik *time series*, analisis *ACF* (*Autocorrelation Function*), dan analisis *PACF* (*Partial Autocorrelation Function*)[3]. Analisa *ACF* dan *PACF* ditujukan untuk dapat mendapatkan informasi hyper parameter  $q$  dan  $p$ . Selain itu analisa *ACF* dan *PACF* juga digunakan pada saat proses *exploratory data analysis (EDA)*, yaitu proses untuk mengetahui karakteristik dari data *time series* yang digunakan untuk peramalan.

Banyak sudah penelitian-penelitian sebelumnya yang sudah dilakukan tentang *forecasting* dengan metode regresi linear dan *ARIMA* seperti Peramalan Tren Pencarian Kata Kunci “Sarung Wadimor” di Indonesia Pada Data Google Trends Menggunakan *Time series Regression with Calender Variation* dan *ARIMA Box-Jenkins*[4], Prediksi Jumlah Kunjungan Wisatawan di Jawa Barat Dengan Model *ARIMA* Dan *SARIMA*

Menggunakan Data Google *Trends* [5], dan Aplikasi Peramalan Penjualan Menggunakan Metode Regresi Linier [6].

Pada penelitian sebelumnya metode *ARIMA* dan regresi linear banyak digunakan dan metode ini cukup efektif karena memberikan hasil *RMSE* dan *MAPE* yang rendah sebesar 0,374 dan 0,280 pada penelitian yang dilakukan oleh Ghebyla Najla Ayuni dan Devi Fitriana, Penerapan Metode Regresi Linear Untuk Prediksi Penjualan Properti pada PT XYZ [7]. Prediksi Jumlah Kunjungan Wisatawan Di Jawa Barat Dengan Model *ARIMA* Dan *SARIMA* Menggunakan Data Google *Trends*, didapatkan hasil *MAPE* sebesar 16,71 [5].

Berdasarkan dari banyaknya penelitian yang telah dilakukan dan melihat kelebihan dari kedua metode tersebut maka, dapat disimpulkan bahwa metode regresi linear dan *ARIMA* dapat digunakan untuk melakukan *forecasting data time series trend keyword SEO* untuk industri *real-estate* yang dalam penelitian ini menggunakan *keyword* perumahan. Penelitian ini akan berfokus untuk membandingkan hasil akurasi regresi linear dan *ARIMA* dalam melakukan *forecasting data time series trend keyword SEO*, penelitian ini akan melakukan pengukuran menggunakan *RMSE* (*root mean squared error*), *MAE* (*mean absolute error*) dan *MAPE* (*mean absolute percentages error*).

Permasalahan yang muncul dalam membuat perbandingan metode regresi linear dan *ARIMA* adalah hasil tingkat akurasi yang beragam (baik dan buruk) dan apakah kedua model tersebut dapat digunakan untuk meramalkan *trend keyword SEO*. Hal ini menjadi faktor yang penting karena jika model yang digunakan tidak memiliki akurasi yang baik tentunya akan mempengaruhi hasil peramalan yang berujung pada tidak dapat digunakannya kedua model tersebut. Untuk itu, pada penelitian ini akan membuat perbandingan metode regresi linear dan *ARIMA* untuk meramalkan *trend keyword SEO* untuk menyelesaikan masalah penelitian kata kunci dalam strategi pemasaran digital di industry *real estate*, dengan pengecekan tingkat *RMSE* dan *MAPE*.

## 1.2. Rumusan Masalah

Berdasarkan latar belakang masalah, maka didapatkan rumusan masalah sebagai berikut:

1. Persaingan *visibilitas website* antara perusahaan *real-estate* dengan menggunakan *trend keyword SEO*.
2. Perbandingan hasil akurasi metode regresi linear dan *ARIMA* untuk meramalkan *trend keyword SEO* dengan kata kunci perumahan.
3. Perbandingan hasil peramalan metode regresi linear dan *ARIMA* untuk meramalkan *trend keyword SEO* dengan kata kunci perumahan.
4. Pengaruh peramalan *trend keyword SEO* dengan metode regresi linear dan *ARIMA* untuk penelitian kata kunci *SEO* pada industri *real-estate*.

## 1.3. Pertanyaan Penelitian

Berdasarkan rumusan masalah yang telah disebutkan di atas, dapat menghadirkan pertanyaan sebagai berikut:

1. Bagaimana pengaruh model regresi linear dan *ARIMA* dalam meramalkan *trend keyword SEO* untuk proses analisa kata kunci *SEO* pada industri *real-estate*?
2. Bagaimana hasil evaluasi model regresi linear dan *ARIMA* dalam meramalkan *trend keyword SEO* untuk proses analisa kata kunci *SEO* pada industri *real-estate* dengan kata kunci perumahan?
3. Bagaimana hasil peramalan dapat digunakan untuk merekomendasikan *trend keyword SEO* dalam menganalisa kata kunci *SEO* pada industri *real-estate* dengan kata kunci perumahan?
4. Bagaimana pengaruh *trend-keyword SEO* dalam persaingan visibilitas website antara Perusahaan *real-estate*?

## 1.4. Batasan Masalah

Batasan masalah pada penelitian ini adalah:

1. Perbandingan metode regresi linear dan *ARIMA* ini hanya terbatas untuk penelitian kata kunci “perumahan”.
2. Hasil akurasi prediksi model yang dibuat dihitung menggunakan *RMSE*, *MAE* dan *MAPE*.

3. Data yang digunakan dalam penelitian ini adalah data *time series* yang didapat dari *google trend* dengan rentang waktu dari tanggal 16 september 2018 sampai dengan tanggal 3 september 2023.
4. *Output* perbandingan model hanya berupa rekomendasi *trend keyword SEO* pada kata kunci “perumahan” yang mengalami kenaikan dimasa depan atau mengalami penurunan.

### 1.5. Tujuan

Berdasarkan rumusan masalah yang ada, maka dapat diketahui tujuan dari penelitian ini adalah:

1. Mengetahui pengaruh model regresi linear dan *ARIMA* dalam meramalkan *trend keyword SEO* untuk proses analisa kata kunci *SEO* pada industri *real-estate*.
2. Menghitung hasil evaluasi model regresi linear dan *ARIMA* dalam meramalkan *trend keyword SEO* untuk proses analisa kata kunci *SEO* pada industri *real-estate* dengan kata kunci perumahan.
3. Mengetahui hasil peramalan dapat digunakan untuk merekomendasikan *trend keyword SEO* dalam menganalisa kata kunci *SEO* pada industri *real-estate* dengan kata kunci perumahan.

### 1.6. Manfaat

Berdasarkan rumusan masalah, tujuan penelitian dan batasan masalah maka dapat diketahui manfaat dari penelitian ini adalah:

1. Penelitian ini diharapkan mampu memberikan wawasan tentang model regresi linear dan *ARIMA* untuk mendapatkan akurasi yang baik dalam meramalkan *trend keyword SEO*.
2. Penelitian ini diharapkan mampu memberikan wawasan tentang penggunaan metode regresi linear dan *ARIMA* untuk meramalkan *trend keyword SEO*.