

BAB I

PENDAHULUAN

1.1 Latar Belakang

Faktor kompresibilitas gas merupakan sebuah rasio dari volume molaritas sebenarnya suatu gas nyata (*real gas*) terhadap volume molaritas gas ideal. Gas nyata adalah sebuah gas yang ada secara alami seperti O₂, H₂, He dan CO₂. Sedangkan gas ideal adalah gas imajiner atau gas hipotetis karena gas ini merupakan gas yang tidak bisa ada secara alami, gas ideal merupakan sebuah konsep gas yang mengikuti hukum gas ideal yaitu sebuah perkiraan perilaku berbagai gas terhadap berbagai kondisi. Hukum gas ideal adalah hukum yang menyatakan *equation of state* gas ideal. Hukum tersebut pertama kali diperkenalkan oleh Emile Clapeyron pada 1834 sebagai hukum yang menggabungkan hukum empiris Boyle, Charles, dan Avogadro (Clapeyron, 1834).

Faktor kompresibilitas gas merupakan sebuah nilai yang menyatakan seberapa jauh sebuah gas nyata melenceng (*deviate*) dari gas ideal. Salah satu contoh penggunaan pencarian nilai simpangan gas nyata terhadap gas ideal ada pada proses *tertiary production* pada *enhanced oil recovery* (EOR). Pada proses tersebut dilakukan proses CO₂ *injection* yaitu proses penyuntikan CO₂ penyuntikan dilakukan dengan menggunakan CO₂ yang mempunyai kemurnian mencapai 95%, CO₂ itu digunakan untuk memicu reaksi CO₂ dengan minyak sehingga menjadikan minyak menjadi mengembang dan membuatnya menjadi lebih ringan yang bertujuan untuk membuat minyak bisa bergerak lebih bebas ke tempat penampungan (Melzer, 2012). Oleh karena itu pencarian nilai faktor kompresibilitas sangat diperlukan untuk menghitung seberapa jauh simpangan gas CO₂ dari gas ideal.

Terdapat beberapa cara untuk mencari nilai faktor kompresibilitas diantaranya yaitu dengan eksperimen, korelasi (Navid, Isazadeh, & Behbahani, 2010) dan dengan menggunakan rumus *equation of state*. Tetapi untuk melakukan eksperimen tersebut memerlukan biaya yang mahal dan waktu yang tidak sedikit. Begitu pula dengan menggunakan rumus hukum gas ideal, hukum gas ideal tidak bisa menghitung secara akurat pada kondisi tekanan tinggi dan suhu rendah. Sehingga diperlukan metode lain yang bisa mencari nilai tersebut. Salah satu cara untuk membantu menyelesaikan permasalahan tersebut adalah dengan menggunakan *machine learning*, yaitu dengan cara membangun model berdasarkan data histori.

Machine learning adalah sebuah ilmu yang mempelajari bagaimana membuat sebuah mesin supaya bisa berfikir dan belajar secara mandiri selayaknya manusia (Samuel, 1959). Pada umumnya terdapat 2 proses utama pada *machine learning* yaitu *training* dan *prediction*. *Training* adalah sebuah proses dimana mesin belajar tentang masalah yang sedang dihadapi dan membuatkan model untuk menyelesaikan permasalahan tersebut, sedangkan *prediction* adalah proses dimana mesin menuangkan hasil prediksi dengan menggunakan data *testing*.

Penelitian sebelumnya tentang prediksi faktor kompresibilitas gas CO₂ telah pernah dilakukan oleh beberapa peneliti seperti oleh (Mohagheghian, Zafarian-Rigaki, Motamedi-Ghahfarrokhi, & Hemmati-Sarapardeh, 2015) yang menggunakan algoritma *artificial neural network* (ANN). Hasil dari penelitian dinyatakan bahwa algoritma ANN bisa memprediksi nilai faktor kompresibilitas dengan akurasi yang baik yaitu hanya gagal menebak 37 dari 2118 data.

Oleh karena itu peneliti akan mencoba menggunakan metode lain untuk memprediksi data tersebut. Metode yang akan penulis gunakan adalah metode *gradient descent* (GD) beserta variannya yang pada penelitian ini metode yang dimaksud adalah *mini-batch gradient decent* (MBGD) (Cotter, Shamir, Srebro, & Sridharan, 2011), *stochastic gradient descent* (SGD) (Bottou, 2010), (Le Roux, Schmidt, & Bach, 2011), dan *stochastic average gradient* (SAG) (Schmidt, Le Roux, & Bach, 2013). Metode GD beserta variannya

termasuk metode untuk menyelesaikan permasalahan *supervised learning* bertipe regresi (*regression*) karena model yang dihasilkan oleh metode ini berupa nilai real.

Terdapat beberapa bahasa pemrograman yang biasa digunakan untuk membuat program *machine learning*, di antaranya *python*, *java*, dan *R*. Berdasarkan survey yang dilakukan KDnuggets (Piatetsky, 2015) menyatakan bahwa bahasa pemrograman *R* merupakan bahasa pemrograman yang banyak digunakan untuk menyelesaikan permasalahan statistik, *data mining*, dan *machine learning*. Bahasa pemrograman *R* menawarkan beberapa kemudahan untuk menyelesaikan berbagai permasalahan matematis seperti menawarkan banyak fungsi bawaan tentang data manipulation seperti transformasi, agregasi, penggabungan dll. Selain itu bahasa pemrograman *R* memiliki dokumentasi yang cukup baik, hal ini terlihat dari terdapatnya *archive* dokumentasi yang tersedia untuk pemrograman ini, seperti CRAN (*Comprehensive R Archive Network*) merupakan sebuah *website* yang berisikan karya para pengembang bahasa pemrograman *R* yang dilamannya terdapat banyak kumpulan karya pengembang *R* dengan berbagai macam topik seperti *machine learning*, diferensial, *time series* dll. Karya yang berada pada *archive* tersebut kebanyakan berupa *R package* yang merupakan kumpulan fungsi, data, dan dokumentasi yang disatukan dalam satu paket direktori. *R package* juga merupakan standarisasi untuk user apabila ingin mengirimkan karyanya ke CRAN.

1.2 Rumusan Masalah

Rumusan masalah yang terdapat pada penelitian ini adalah:

1. Bagaimana implementasi metode GD, MBGD, SGD, dan SAG pada bahasa pemrograman *R*?
2. Bagaimana pembuatan model perhitungan faktor kompresibilitas gas dengan menggunakan metode pada persamaan 1.
3. Berapa tingkat akurasi model yang dibuat oleh metode pada poin 1 untuk studi kasus perhitungan faktor kompresibilitas?

1.3 Batasan Masalah

Batasan masalah yang terdapat pada penelitian ini adalah:

1. Data yang digunakan hanya data tentang perilaku gas CO₂ pada temperatur dan tekanan tertentu.
2. Data yang digunakan merupakan data yang didapat dari hasil penelitian yang dilakukan (Kennedy, 1954).

1.4 Tujuan Penelitian

Tujuan penelitian ini adalah:

1. Mengimplementasikan metode *Gradient Descent* beserta variasi dari metode *Gradient Descent* dalam hal ini yaitu *Stochastic Gradient Descent*, *Mini Batch Gradient Descent*, dan *Stochastic Average Gradient Descent* pada bahasa pemrograman R.
2. Membuat model untuk prediksi faktor kompresibilitas dengan menggunakan yang disebutkan pada poin 1.
3. Mengukur akurasi model yang dibuat dari metode metode yang dipakai dengan cara menghitung seberapa besar kesalahan yang didapat pada setiap simulasi.

1.5 Manfaat Penelitian

Manfaat penelitian ini adalah:

1. Dibuatnya *software* yang bisa memprediksi contoh kasus klasifikasi berbentuk regresi dengan menggunakan 4 metode yaitu: *gradient descent*, *stochastic gradient descent*, *mini-batch gradient descent* dan *stochastic average gradient*.
2. Dengan dilakukan penelitian ini bisa didapatkan metode apa yang merupakan metode terbaik untuk memprediksi nilai faktor kompresibilitas gas CO₂ serta melihat seberapa akurasi dari masing masing metode yang dipakai untuk menyelesaikan permasalahan tersebut.

1.6 Sistematika Penulisan

Sistematika penulisan penelitian ini adalah:

BAB I PENDAHULUAN

Bagian ini merupakan gambaran umum dari penelitian. Meliputi latar belakang, rumusan masalah, tujuan penelitian, batasan masalah, metodologi penelitian dan sistematika penulisan.

BAB II KAJIAN PUSTAKA

Bagian ini merupakan penjelasan secara mendetail tentang materi yang diperlukan pada penelitian ini. Dalam hal ini penjelasan mendalam tentang materi Bahasa pemrograman R, *Machine Learning*, Faktor kompresibilitas gas, GD dan variannya.

BAB III METODOLOGI PENELITIAN

Bagian ini merupakan penjelasan akan proses penelitian yang dilakukan mulai dari alat dan bahan yang digunakan pada penelitian ini, desain penelitian, serta proses apa saja yang dilakukan untuk menyelesaikan penelitian ini.

BAB IV HASIL PENELITIAN

Bagian ini merupakan bagian yang mencantumkan hasil dari penelitian yang telah dilakukan dengan menggunakan metode yang digunakan pada penelitian ini.

BAB V KESIMPULAN DAN SARAN

Bagian ini berisi tentang hasil dari penelitian berupa kesimpulan yang merupakan jawaban dari masalah dalam penelitian, serta berisi saran untuk penelitian selanjutnya.