

# BAB I

## PENDAHULUAN

### 1.1 Latar Belakang

Air merupakan komponen yang sangat penting bagi makhluk hidup terutama manusia. Air ini banyak dimanfaatkan manusia untuk kebutuhannya sehari-hari misalnya untuk konsumsi rumah tangga, kebutuhan industri, dan lain sebagainya. Dalam memenuhi kebutuhan air tersebut diperlukan air yang bersih dimana air itu tidak tercemar dari berbagai hal yang dapat merugikan kesehatan tubuh maupun lingkungan. Permasalahan air di seluruh dunia mencapai titik krisis, didukung oleh data dari Perserikatan Bangsa Bangsa (PBB) pada tahun 2023, yang menyebutkan bahwa 2,2 miliar orang atau seperempat populasi dunia kekurangan air untuk konsumsi sehari-hari [2]. Adapun menurut laporan Bappenas, ketersediaan air di sebagian besar wilayah Pulau Jawa dan Bali saat ini tergolong langka hingga kritis [3]. Karena kebutuhan air yang sangat penting itu maka diperlukan adanya penanganan terkait masalah ketersediaan air ini.

Kota DKI Jakarta adalah salah satu kota besar sekaligus ibu kota di Indonesia, hal ini berdampak pada ketersediaan air yang dibutuhkan oleh masyarakat. Untuk memenuhi kebutuhan air masyarakat DKI Jakarta, pemerintah menyediakan pusat air bersih yaitu Perusahaan Daerah Air

Minum (PAM JAYA) untuk masyarakat DKI Jakarta. Dalam memenuhi kebutuhan air ini, pemerintah harus memperhatikan kuantitas air yang harus didistribusikan kepada masyarakat karena adanya faktor persoalan akan kebutuhan air ini. Salah satu persoalannya yaitu jika jumlah air yang diproduksi dan disalurkan lebih besar dari permintaan air, perusahaan akan menghadapi masalah pemborosan air. Sebaliknya, jika jumlah air yang diproduksi dan disalurkan lebih sedikit atau tidak memenuhi kebutuhan konsumen, akan terjadi kekurangan volume air, yang akan merugikan konsumen [4]. Oleh karena itu, diperlukan prediksi yang tepat dalam memperkirakan jumlah produksi air yang akan diproduksi untuk didistribusikan kepada masyarakat setempat.

Data jumlah produksi air ini merupakan data deret waktu. Ada banyak metode yang dapat digunakan untuk memprediksi data deret waktu ini. Salah satunya yaitu *Autoregressive Integrated Moving Average (ARIMA)*. *ARIMA* adalah metode peramalan yang biasanya sangat baik digunakan untuk melakukan peramalan jangka pendek. Akan tetapi, metode *ARIMA* memiliki keterbatasan dimana metode ini membutuhkan data historis dalam jumlah yang banyak agar menghasilkan prediksi yang akurat. Karena keterbatasan itulah, diciptakan sebuah metode *deep learning* yang membantu permasalahan data jangka pendek ini seperti metode *Long Short-Term Memory (LSTM)* [5].

Penelitian Siami-Namini dkk [6] telah membandingkan kinerja antara metode *ARIMA* dan metode *deep learning*, khususnya *LSTM*. Pada metode *ARIMA* dan *LSTM* diperoleh *RMSE* masing-masing sebesar 511.481

dan 64.213. Hasil penelitian menunjukkan bahwa algoritma berbasis *LSTM* meningkatkan akurasi prediksi sebesar 85% secara rata-rata dibandingkan dengan *ARIMA*. Metode *ARIMA* memerlukan penyesuaian parameter ( $p$ ,  $d$ ,  $q$ ) yang signifikan dimana parameter ini akan mempengaruhi hasil prediksi. Jika parameter ( $p$ ,  $d$ ,  $q$ ) yang digunakan tidak optimal, maka prediksi yang dihasilkan dapat menjauhi data aktual. Di sisi lain, *LSTM* berbasis *deep learning* mencapai prediksi optimal dengan melakukan iterasi berulang hingga ditemukan hasil prediksi dan model yang paling sesuai dengan data.

Pada tahun 1980, Granger dan Joyeux memperkenalkan model *ARFIMA*, sebuah kemajuan dari model *ARIMA* yang mampu memprediksi data deret waktu berpola memori jangka pendek dan panjang. Model nonlinear melibatkan penggunaan pemrosesan numerik atau iteratif pada data yang besar, sehingga memungkinkan estimasi yang akurat terhadap nilai sebenarnya [7].

Metode lain yang dapat digunakan untuk mengolah data deret waktu yaitu metode *RNN* (*Recurrent Neural Network*). Metode *Recurrent Neural Network* (*RNN*) adalah *neural network* berulang yang didesain khusus untuk mengolah data berurutan. Namun, *RNN* memiliki kekurangan dalam memprediksi data yang bersifat jangka panjang. Untuk menangani masalah tersebut, digunakanlah model *RNN* tipe *Long Short Term Memory* yang dibuat untuk meminimalisir error dan menghindari masalah ketergantungan jangka panjang pada *Recurrent Neural Network* (*RNN*) [8].

Penelitian yang dilakukan oleh Yanto dkk[9], hasil pengujian yang

telah dilakukan dengan algoritma *backpropagation* diperoleh nilai akurasi pada peramalan jumlah produksi air sebesar 99,78% dan nilai rata-rata kesalahannya (*MAPE*) yang didapat sebesar 0.23%. Prediksi ini sudah sangat bagus karena nilai akurasinya yang mendekati 100% dan juga nilai rata-rata kesalahannya yang sangat kecil. Namun, pemilihan pola arsitektur dari prediksi ini harus lebih diperhatikan karena akan mempengaruhi hasil prediksi.

Menurut Lattifia dkk [10] , dalam memprediksikan cuaca digunakan metode *LSTM* yang dapat mendeteksi data yang akan disimpan dan data yang tidak digunakan untuk dipangkas. Hasil prediksi yang diperoleh yaitu nilai akurasi terbaik rata-rata menggunakan *batch size* 50 serta *epoch* 100 dan nilai *RMSE* dan *MAPE* sebesar 1,7444 dan 1,9499%.

Berdasarkan dari penelitian-penelitian tersebut, terlihat bahwa metode *Long Short Term Memory (LSTM)* mengantisipasi *epoch* dan ukuran *batch*. Selain itu, kinerja dari metode *LSTM* yang sangat baik membuat penelitian tentang metode ini menjadi sangat luas. Namun, sedikit sekali penelitian yang menerapkan model-model *LSTM* untuk memprediksi parameter kuantitas air (dalam hal ini jumlah produksi air) [11]. Oleh karena itu, dilakukan prediksi dengan menggunakan metode ini terhadap jumlah produksi air PAM terkhususnya di Provinsi DKI Jakarta.

## 1.2 Rumusan Masalah

Berdasarkan latar belakang di atas, maka permasalahan yang akan dikaji pada penelitian tugas akhir ini adalah:

1. Bagaimana hasil prediksi jumlah produksi air PAM yang ada di Provinsi DKI Jakarta dengan menggunakan metode *Long Short Therm Memory*?
2. Bagaimana tingkat akurasi dari hasil prediksi jumlah produksi air PAM yang ada di Provinsi DKI Jakarta dengan menggunakan metode *Long Short Therm Memory*?

## 1.3 Batasan Masalah

Batasan masalah dari penelitian ini yaitu prediksi yang dilakukan dalam penelitian ini terbatas pada memperkirakan jumlah produksi air PAM di masa depan. Penelitian tidak mempertimbangkan variabel lain yang mungkin memengaruhi produksi air, seperti perubahan iklim atau pertumbuhan populasi.

## 1.4 Tujuan Penulisan

Adapun tujuan penulisan tugas akhir ini adalah :

1. Untuk memperoleh hasil prediksi jumlah produksi air PAM yang ada di Provinsi DKI Jakarta dengan menggunakan metode *Long Short Therm Memory*.

2. Untuk mengetahui tingkat akurasi dari hasil prediksi jumlah produksi air PAM yang ada di Provinsi DKI Jakarta dengan menggunakan metode *Long Short Term Memory*.

## 1.5 Sistematika Penulisan

Sistematika penulisan tugas akhir ini terdiri dari lima bab. Bab I memuat latar belakang, rumusan masalah, batasan masalah, tujuan penulisan, dan sistematika penulisan. Bab II memuat teori dasar dan materi pendukung yang berkaitan dengan topik penelitian tugas akhir ini. Selanjutnya, Bab III berisikan langkah-langkah yang akan dilakukan untuk menyelesaikan masalah yang ada pada rumusan masalah dalam penelitian tugas akhir. BAB IV yang berisi hasil dan pembahasan dari penelitian tugas akhir. Langkah akhir yaitu, BAB V berisikan kesimpulan dari hasil penelitian dan saran untuk penelitian selanjutnya.

