

# BAB I PENDAHULUAN

## 1.1 Latar Belakang

Peramalan merupakan suatu kegiatan yang digunakan untuk memperkirakan apa yang akan terjadi di masa yang akan datang dengan memanfaatkan informasi kejadian atau peristiwa yang ada pada masa kini dan masa lalu. Tujuannya agar dapat melakukan perencanaan kegiatan atau peristiwa di masa yang akan datang. Berdasarkan sifatnya peramalan dibagi menjadi dua, yaitu peramalan kualitatif dan peramalan kuantitatif. Peramalan kualitatif merupakan peramalan yang dilakukan berdasarkan pendapat para ahli, dan peramalan kuantitatif merupakan peramalan yang dilakukan berdasarkan data dan informasi yang sudah ada sebelumnya (Herawati dkk., 2016). Berdasarkan data historis, Restu Wardhani & Manuel Pereira (2010) menyatakan bahwa peramalan kuantitatif terbagi menjadi dua, yaitu metode runtun waktu (*time series*) dan metode kausal.

Salah satu metode runtun waktu yang dapat digunakan untuk melakukan peramalan adalah metode *Autoregressive Integrated Moving Average* (ARIMA), yang dikembangkan oleh Box Jenkins pada awal 1970-an. Ketepatan peramalan dengan menggunakan metode ARIMA sangat baik untuk jangka waktu pendek. Pada penerapannya, metode ARIMA hanya menggunakan variabel dependen dan mengabaikan variabel independen untuk menghasilkan peramalan jangka pendek yang akurat (Salwa dkk., 2018).

Pada data runtun waktu terdapat empat jenis pola data, yaitu pola horizontal, pola siklis, pola *trend*, dan pola musiman (Makridakis dkk. 1999). Penting untuk memperhatikan pola data runtun waktu yang dimiliki agar metode peramalan yang digunakan sesuai dengan karakteristik yang ada. Salah satu metode pengembangan dari ARIMA yang cocok untuk meramalkan pola data musiman adalah *Seasonal Autoregressive Integrated Moving Average* (SARIMA). Pola musiman yang dimaksud merupakan pola data yang memiliki kecenderungan untuk mengulangi bentuk pola data yang terjadi dalam kurun waktu satu tahun, seperti semester, triwulan, bulanan, mingguan, atau harian. Namun, metode ARIMA maupun SARIMA ini memiliki kekurangan yaitu mengabaikan kemungkinan hubungan non

linear, dan homokedastisitas residual (Adnyani & Subanar, 2015). Sedangkan pada kenyataannya, tidak semua data memiliki karakteristik yang linear. Maka dari itu, diperlukan metode lain untuk dapat melakukan peramalan terhadap data yang non linear. Salah satu metodenya adalah *Long Short Term Memory* (LSTM).

Berbeda dengan metode SARIMA, LSTM merupakan salah satu jaringan syaraf tiruan yang dikembangkan dari algoritma *Reccurent Neural Network* yang dapat digunakan dalam meramalkan data runtun waktu non linear, di mana LSTM mampu mengekstraksi informasi dari data jangka panjang, runtun waktu atau sequental (Julian & Pribadi, 2021). LSTM dikenal dengan kemampuannya untuk memodelkan dan memprediksi runtun waktu, tetapi penggunaan LSTM membutuhkan data runtun waktu yang lama (Peirano dkk., 2021). Namun, salah satu permasalahan yang muncul pada data runtun waktu jangka panjang yaitu terjadinya *error* kecil yang dapat menyebar sehingga menyebabkan kegagalan model konvergen (Petneházi, 2018). Kesesuaian LSTM untuk meramalkan runtun waktu yang lama dibentuk sedemikian rupa sehingga dapat mengingat informasi yang tidak dapat dilakukan oleh model *Neural Network* (NN) sederhana, sementara pada saat yang sama menggunakan sel memori jangka panjang yang dapat menyelesaikan masalah *vanishing gradient* dari RNN yang dapat mempengaruhi kinerja hasil dalam peramalan (Parasyris dkk., 2022). Kelebihan lain yang dimiliki LSTM, yaitu LSTM sangat baik dalam mengingat informasi dalam jangka waktu yang lama (Kumar dkk., 2018).

Indonesia merupakan negara dengan iklim tropis di mana hanya mempunyai dua musim dalam setahun, yakni musim kemarau dan musim hujan. Oleh karena itu, Indonesia memiliki penguapan dan intensitas curah hujan yang cukup tinggi (Adib Azka dkk., 2018). Rizki dkk., (2020) menyatakan bahwa tiap daerah memiliki intensitas curah hujan yang berbeda-beda yang disebabkan oleh beberapa faktor yakni garis lintang, ketinggian daerah, jarak dari sumber air, arah angin, suhu tanah, dan luas daratan. Pada kenyataannya cuaca dapat tiba-tiba berubah tidak sesuai dengan musimnya.

Bogor merupakan salah satu kota besar di Indonesia yang terletak di dataran tinggi yang sering terjadi hujan. Kepala Stasiun Badan Meteorologi, Klimatologi, dan Geofisika Darmaga, Budi Suhardi (2021) mengatakan jika rata-rata curah hujan

melebihi 50 mm, maka disebut dengan musim hujan. Sedangkan, dalam hitungan sepuluh hari, intensitas curah hujan yang terjadi di Bogor rata-rata di atas 50 mm, yang artinya banyak terjadi hujan. Curah hujan tinggi yang terjadi di Bogor menyebabkan ancaman terjadinya bencana alam seperti banjir dan tanah longsor. Kondisi curah hujan yang tinggi juga memiliki dampak yang besar pada kehidupan sehari-hari seperti transportasi, pertanian, dan industri lainnya. Pada bidang transportasi dapat mempengaruhi kelancaran jalur transportasi baik darat, laut maupun udara. Pada bidang pertanian, dapat mempengaruhi hasil panen. Terjadinya curah hujan tidak dapat ditentukan secara pasti, namun dapat diramal atau diprediksi (Oktaviani & Afdal, 2013). Dengan memanfaatkan informasi mengenai curah hujan di masa lampau, maka dapat memprediksi curah hujan yang akan terjadi pada masa yang akan datang. Sehingga pemerintah dan masyarakat dapat menyiapkan langkah selanjutnya untuk mengantisipasi dan menghadapi jika terjadi curah hujan yang tinggi.

Penggabungan atau *hybrid* pada pemodelan data linear dan non linear bertujuan untuk dapat mengatasi data dengan pola linear dan non linear. Peramalan *hybrid* artinya melakukan kombinasi atau penggabungan antar beberapa model. Tujuan dilakukannya *hybrid* antar model karena ingin meminimalkan kekurangan yang ada pada masing-masing metode dan memanfaatkan kelebihanannya (Muslim, 2017). Tujuan lain dengan dilakukannya *hybrid* yaitu, untuk melakukan peramalan yang lebih akurat dibanding jika menggunakan metode secara individu (Gunaryati dkk., 2019). Beberapa penelitian sebelumnya yang telah menggunakan model *hybrid* SARIMA-LSTM di antaranya adalah peramalan data inflasi oleh Peirano dkk. (2021), kemudian peramalan data kedatangan turis oleh Wu dkk. (2021), dan *hybrid* ARIMA-LSTM dalam meramalkan harga saham oleh (Wibawa, 2021).

Berdasarkan uraian latar belakang di atas, penulis tertarik untuk meneliti model *hybrid* dalam runtun waktu yang berfokus pada penggabungan model SARIMA dengan LSTM untuk meramalkan data curah hujan. Oleh karena itu, penelitian ini diberi judul “Penerapan *Hybrid Seasonal Autoregressive Integrated Moving Average – Long Short Term Memory* (SARIMA-LSTM) dalam Meramalkan Curah Hujan di Bogor”.

## 1.2 Rumusan Masalah

Berdasarkan uraian latar belakang di atas, maka rumusan masalah yang akan dibahas pada penelitian ini adalah sebagai berikut:

1. Bagaimana model *Hybrid* SARIMA-LSTM yang terbaik untuk melakukan peramalan curah hujan di Bogor?
2. Bagaimana hasil peramalan curah hujan di Bogor dengan menggunakan model *Hybrid* SARIMA-LSTM yang terbaik?

## 1.3 Tujuan Penelitian

Berdasarkan rumusan masalah yang telah dipaparkan, adapun tujuan dari penelitian ini adalah:

1. Untuk memperoleh model *Hybrid* SARIMA-LSTM yang terbaik untuk melakukan peramalan curah hujan di Bogor.
2. Untuk memperoleh hasil peramalan curah hujan di Bogor dengan menggunakan model *Hybrid* SARIMA-LSTM yang terbaik.

## 1.4 Manfaat Penelitian

Adapun manfaat yang dapat diambil dari penelitian ini adalah sebagai berikut:

1. Manfaat Teoritis:  
Hasil penelitian ini dapat memberikan wawasan baru mengenai model peramalan runtun waktu, khususnya model *Hybrid* SARIMA-LSTM.
2. Manfaat Praktis:
  - a. Peneliti dapat mengaplikasikan ilmu matematika khususnya dalam bidang statistika yang dipelajari di perkuliahan pada permasalahan kehidupan sehari-hari.
  - b. Penelitian ini dapat digunakan sebagai referensi dalam ilmu matematika khususnya pada bidang statistika terkait penerapan model *Hybrid* SARIMA-LSTM dalam peramalan curah hujan.
  - c. Memberikan tambahan informasi sebagai bahan pertimbangan bagi Badan Meteorologi, Klimatologi, dan Geofisika (BMKG) untuk peramalan curah hujan di masa yang akan datang.

### 1.5 Batasan Masalah

Untuk menghindari pembahasan yang terlalu luas pada penelitian yang dilakukan maka diperlukan batasan-batasan masalah sebagai berikut:

1. Data yang digunakan yaitu data sekunder yang diperoleh dari *website* Badan Meteorologi, Klimatologi, dan Geofisika (BMKG) sejak Januari 1985 sampai dengan Desember 2021.
2. Data yang digunakan merupakan data jumlah curah hujan perbulan yang diperoleh dari Stasiun Meteorologi Citeko Bogor.
3. Pengolahan data pada penelitian skripsi ini menggunakan bantuan *software RStudio, Python, dan Microsoft Excel*.