

Bab I Pendahuluan

I.1 Latar Belakang

Demam Berdarah Dengue (DBD) merupakan penyakit menular yang disebabkan oleh virus *dengue* dan ditularkan melalui gigitan nyamuk *Aedes aegypti*. DBD merupakan penyakit endemik yang menyerang daerah yang memiliki iklim tropis dan subtropis (WHO Regional Office for South-East Asia, 2011). Penyakit ini berpotensi menyerang 1,8 miliar orang yang tinggal di negara-negara yang berada di wilayah Asia Tenggara, Afrika, Amerika, bagian Mediterania Timur, dan bagian barat Pasifik. World Health Organization (WHO) memperkirakan bahwa setiap tahunnya terjadi 50 juta hingga 100 juta kasus DBD (World Health Organization, 2009).

Di Indonesia tercatat pada tahun 2015 kasus DBD sebanyak 129.650 kasus dan 1.071 orang dinyatakan meninggal dunia (Kementerian Kesehatan Republik Indonesia, 2015). Pada tabel I.1 di bawah ini menunjukkan kasus DBD yang terjadi Provinsi DKI Jakarta tahun 2018.

Tabel I.1 Persebaran penyakit DBD di DKI Jakarta tahun 2018

No.	Daerah	Kasus
1	Jakarta Barat	947
2	Jakarta Timur	876
3	Jakarta Selatan	738
4	Jakarta Utara	278
5	Jakarta Pusat	118
6	Kepulauan Seribu	10

(Sumber: Dinas Kesehatan DKI Jakarta)

Provinsi DKI Jakarta memiliki *Incident Rate* (IR) kasus DBD sebesar 48,67 kasus per 100.000 penduduk pada tahun 2015 (Pusat Data dan Informasi Kementerian Kesehatan Republik Indonesia, 2016). IR adalah tingkat terjadinya suatu kasus per 100.000 penduduk. Semakin tinggi IR berarti kasus terjadi semakin banyak.

Penyakit DBD merupakan penyakit endemik yang selalu dipantau oleh Dinas Kesehatan DKI Jakarta untuk menekan kasus DBD menjadi kejadian luar biasa (KLB). Oleh karena itu perlu dilakukannya *forecasting* yang dapat membantu Dinas Kesehatan DKI Jakarta dalam menekan persebaran penyakit DBD. Untuk melakukan *forecasting* kasus demam berdarah di DKI Jakarta dapat dilakukan dengan *data mining*. *Data mining* adalah sebuah proses menemukan wawasan atau suatu pola tertentu dalam sekumpulan data yang berjumlah sangat banyak (Han and Kamber, 2006). Salah satu algoritma *data mining* yang dapat digunakan untuk *forecasting* peningkatan kasus DBD adalah algoritma regresi linear. Regresi linear merupakan suatu algoritma untuk menemukan hubungan antara suatu variabel dengan variabel yang lain (Cichosz, 2015). Regresi linear juga dapat digunakan sebagai prediksi terhadap suatu nilai yang akan datang dengan data-data historis (*time series forecasting*). Penelitian yang telah dilakukan sebelumnya untuk memprediksi peningkatan kasus DBD pada Provinsi DIY Yogyakarta dan didapatkan nilai eror menggunakan regresi linear sebesar 0,0212 untuk *mean absolute error* (MAE) dan 0,0453 untuk *root mean square error* (RMSE) (Qomaruddin Munir and Winarko, 2015).

Agar Dinas Kesehatan DKI Jakarta dapat melakukan tindakan yang tepat untuk menekan persebaran DBD maka diperlukannya algoritma yang akurat untuk memprediksi kasus DBD yang akan terjadi. Akurasi prediksi dari algoritma dikatakan akurat apabila nilai dari MAE dan RMSE semakin mendekati nol (Tanwar and Kakkar, 2017). Algoritma lain yang dapat digunakan untuk melakukan *neural network*. Pada tugas akhir ini akan membandingkan akurasi algoritma regresi linear dengan algoritma *neural network*. Alasan dibandingkannya algoritma *neural network* dan regresi linear karena pada penelitian sebelumnya kedua algoritma tersebut akurat dalam memprediksi kasus DBD yaitu pada prediksi kasus

DBD di Kabupaten Sleman menggunakan *neural network* dan didapatkan nilai *Mean Square Error* sebesar 0,0099714 (Wahab, 2016) dan prediksi DBD menggunakan regresi linear di DIY Yogyakarta dan didapatkan nilai *Mean Absolute Error* sebesar 0,0212 dan *Root Mean Square Error* sebesar 0,0453 (Qomaruddin Munir and Winarko, 2015). Untuk mengukur akurasi dari kedua algoritma tersebut digunakan metode *Mean Absolute Error* (MAE), *Mean Absolute Percentage Error* (MAPE), dan *Root Mean Square Error* (RMSE). Semakin kecil nilai MAE, MAPE, dan RMSE berarti semakin akurat hasil dari algoritma tersebut. Mengingat bahwa diperlukannya algoritma yang lebih akurat dalam memprediksi peningkatan kasus DBD pada DKI Jakarta.

Pada penelitian ini akan menggunakan variabel cuaca yaitu temperatur, kelembapan, dan curah hujan karena variabel tersebut berpengaruh terhadap peningkatan kasus DBD (Yudhastuti and Vidiyani, 2005; Choi *et al.*, 2016). Penelitian sebelumnya di Subang Jaya, Malaysia menggunakan ARIMA untuk membuat model prediksi DBD dengan variabel temperatur, kelembapan, dan curah hujan (Dom *et al.*, 2013). Penelitian lainnya yang dilakukan di Daerah Ibu Kota Filipina menggunakan SARIMA untuk memprediksi kasus DBD dengan variabel kelembapan, curah hujan, rata-rata temperatur, minimum temperatur, dan maksimum temperatur (Cruz *et al.*, 2012). Penelitian lain yang dilakukan di Cagayan de Oro, Filipina untuk memprediksi DBD menggunakan regresi linear dengan variabel temperatur, kelembapan, dan curah hujan (Lubos, 2016).

Hasil dari penelitian ini dapat digunakan oleh Dinas Kesehatan DKI Jakarta dalam memprediksi terjadinya kasus DBD yang akan terjadi selama satu tahun sehingga Dinas Kesehatan DKI Jakarta dapat melakukan penanggulangan kasus DBD sehingga kasus DBD dapat mengalami penurunan. Penelitian ini menggunakan bahasa pemrograman R dalam menentukan perbandingan akurasi antara regresi linear dan *neural network*.

I.2 Perumusan Masalah

Permasalahan yang ada dalam tugas akhir ini adalah bagaimana akurasi algoritma *neural network* dan regresi linear dalam memprediksi kasus DBD di DKI Jakarta

I.3 Tujuan Penelitian

Dari permasalahan yang ada tujuan diadakannya penelitian adalah sebagai berikut:

1. Menggunakan algoritma regresi linear dan *neural network* dalam memprediksi peningkatan penyakit DBD di DKI Jakarta
2. Membandingkan akurasi regresi linear dan *neural network* dalam memprediksi kasus DBD di DKI Jakarta

I.4 Manfaat Penelitian

Manfaat penelitian ini adalah:

1. Bagi penulis penelitian ini bermanfaat untuk memberikan gambaran implementasi model regresi linear dan *neural network* dalam membuat analisis prediksi
2. Bagi Dinas Kesehatan DKI Jakarta penelitian ini bermanfaat untuk memberikan informasi prediksi peningkatan penyakit DBD untuk dilakukan upaya pencegahan dan penanggulangan agar penyakit DBD tidak menyebar dan menimbulkan KLB.

I.5 Ruang Lingkup Penelitian

Lingkup penelitian ini adalah:

1. Variabel untuk melakukan prediksi adalah temperatur, kelembapan, dan curah hujan di Provinsi DKI Jakarta
2. Dataset DBD yang digunakan merupakan dataset pemantauan penyakit endemik yang diperoleh dari Dinas Kesehatan DKI Jakarta
3. Dataset temperatur, kelembapan, dan curah hujan yang digunakan diperoleh dari Dinas Lingkungan DKI Jakarta
4. Perbandingan regresi linear dengan *neural network* dilakukan di Provinsi DKI Jakarta dengan data dari 4 kota administratif yaitu Jakarta Pusat, Jakarta Timur, Jakarta Selatan dan Jakarta Utara

5. Perbandingan regresi linear dengan *neural network* dilakukan di 4 kota administratif yaitu Jakarta Pusat, Jakarta Timur, Jakarta Selatan dan Jakarta Utara