

BAB 1. PENDAHULUAN

1.1 Latar Belakang

Demam Berdarah *Dengue* atau disingkat DBD adalah penyakit yang disebabkan oleh virus *dengue* dengan perantara nyamuk *Aedes Aegypti* (Nugroho, 2010). Penyakit ini ditemukan di daerah tropis, subtropis dan menjangkit luas di Asia Tenggara, termasuk Indonesia. Terdapat 4 jenis virus *dengue* yaitu DEN 1, DEN 2, DEN 3 dan DEN 4, masing-masing dapat menyebabkan demam berdarah, baik ringan maupun fatal. Berdasarkan survei Dinas Kesehatan (Dinkes) Kabupaten Jember mencatat selama Januari 2015 dari 300 kasus pasien DBD, 7 orang diantaranya meninggal dunia. Pekan pertama bulan Februari 2015 sudah terjadi 42 kasus. Hal ini menunjukkan sedemikian cepat penyebaran wabah DBD.

Dalam penelitiannya, Bitari menunjukkan bahwa metode fuzzy dapat menjadi solusi untuk melihat potensi penyebaran penyakit DBD di Kabupaten Jember. Parameter-parameter yang digunakan antara lain, curah hujan (CH), hari hujan (HH), angka bebas jentik (ABJ) dan *house index* (HI). Data yang digunakan mulai tahun 2009 sampai tahun 2012 dari 31 kecamatan. Akurasi penelitian tersebut mencapai 75% dari 12 data uji yang diambil acak. Kekurangan penelitian tersebut tidak membahas cara-cara untuk melengkapi data masukan yang kosong (*missing inputation*), faktanya data survei atau data di lapangan sering kali tidak lengkap (Bitari, 2014).

Sementara Keza dalam tugas akhirnya melakukan identifikasi penyakit tanaman kakao untuk dapat membantu masyarakat umum khususnya para petani kakao dalam mengidentifikasi dan menangani tanaman kakao yang terserang penyakit. Program aplikasi identifikasi penyakit tanaman kakao ini dapat sebagai alat bantu untuk pembelajaran tentang gejala-gejala penyakit tanaman kakao. Variabel-variabel yang diperhatikan antara lain cuaca, temperatur, kecepatan angin, dan berolah-raga. Metode yang digunakan dalam mengklasifikasikan penyakit tanaman kakao adalah metode *Naïve Bayes* yang memiliki akurasi 90.9091% dan eror sebesar 9.0909% (Keza, 2014). Dan Sukma dalam

penelitiannya mengklasifikasikan calon pendonor menggunakan metode *Naïve Bayes Classifier* diharapkan membantu kinerja PMI agar lebih cepat dan tepat. Penggunaan metode *Naïve Bayes* sebagai metode utama dalam pengklasifikasian karena metode ini merupakan pengklasifikasi paling sederhana dari model pengklasifikasian dengan peluang. Data yang digunakan dalam penelitian terdapat 2 atribut data yaitu diskrit dan kontinyu. Variabel-variabel yang digunakan yaitu hb, tensi atas, tensi bawah, umur, berat badan, jenis kelamin, riwayat penyakit menular, interval donor (Sukma, 2013).

Dari penelitian yang telah ada terdapat banyak sekali kasus hilangnya nilai pada dataset atau ketiadaan nilai pada data untuk atribut tertentu. Permasalahan hilangnya nilai pada data ini lebih sering disebut dengan *missing imputation*. Penyebab terjadinya *missing imputation* adalah tidak adanya respon terhadap unit maupun item, hal ini merupakan permasalahan yang terjadi pada sebagian survei dan menjadikan hasil dari penelitian kurang baik. Permasalahan *missing imputation* juga dijumpai pada survei tahunan perusahaan industri besar dan sedang yang merupakan salah satu survei rutinyang dilakukan Badan Pusat Statistik. Padahal data yang dihasilkan survey tersebut memiliki manfaat yang penting untuk dianalisis.

Ketiga penelitian diatas mendukung alasan untuk dilakukan penelitian tentang *missing imputation* dataset persebaran demam berdarah *denguedi* Kabupaten Jember dengan metode *Naïve Bayes*. Pada penelitian yang akan dilakukan data yang digunakan adalah data faktor – faktor penyebaran penyakit terhadap demam berdarah *dengue* yaitu curah hujan (CH), hari hujan (HH), angka bebas jentik (ABJ) dan *house indeks* (HI). Data yang digunakan mulai tahun 2009 sampai dengan tahun 2012 dari 31 kecamatan di Kabupaten Jember.

1.2 Perumusan Masalah

Berdasarkan latar belakang diatas, maka permasalahan yang dapat diambil dalam tugas akhir ini adalah sebagai berikut :

- a. Bagaimana memprediksi data yang kosong mengacu pada data-data yang lain?

- b. Bagaimana metode *Naïve Bayes* mendapatkan nilai untuk data yang kosong?

1.3 Batasan Masalah

Mengacu pada masalah yang telah dirumuskan, maka batasan masalah yang akan dibahas dalam pembuatan Tugas Akhir (TA) ini adalah sebagai berikut:

- a. Data yang digunakan adalah data tahun 2009 sampai tahun 2012.
- b. Parameter yang digunakan adalah curah hujan (CH), hari hujan (HH), angka bebas jentik (ABJ), dan *house index* (HI).
- c. Area yang dijadikan objek penelitian adalah 31 kecamatan yang ada di Kabupaten Jember.
- d. Satuan terkecil dalam tabulasi pengelompokan wilayah adalah kecamatan.

1.4 Tujuan

Tujuan dari tugas akhir ini adalah menghasilkan suatu aplikasi untuk prediksi *missing imputation* pada data persebaran demam berdarah *dengue* berdasarkan metode *naïve bayes* dengan *input* yang berupa data inputan parameter yang telah ditentukan, sedangkan *output* berupa hasil prediksi yang diharapkan mampu untuk :

- a. Merancang suatu aplikasi yang dapat melengkapi data kosong/*missing imputation* dengan tingkat akurasi yang tinggi.
- b. Mengimplementasikan metode *Naïve Bayes* pada data nyata guna memprediksi data.

1.5 Manfaat

Adapun manfaat yang diharapkan dari tugas akhir ini antara lain sebagai berikut :

- a. Menghasilkan suatu aplikasi khusus yang dapat melengkapi data kosong/*missing imputation* dengan tingkat akurasi yang tinggi.
- b. Dapat memprediksi suatu data menggunakan metode *Naïve Bayes*.