

BAB I

PENDAHULUAN

Pada bagian ini dikemukakan hal yang melatar belakangi pentingnya dilakukan penelitian, rumusan masalah, dan batasan permasalahan yang dilakukan dalam penelitian, tujuan dan relevansi atau manfaat penelitian terhadap perkembangan solusi dari permasalahan yang diangkat, serta sistematika penulisan dalam laporan Tugas Akhir ini.

1.1 Latar Belakang

RISKESDAS (Riset Kesehatan Dasar) adalah sebuah penelitian kuantitatif yang dilakukan oleh badan litbang Kementerian Kesehatan secara berkala dalam kurun waktu tiga tahunan. Badan LITBANG (Penelitian dan Pengembangan) Kementrian Kesehatan melakukan penelitian tentang tingkat kesehatan masyarakat diseluruh propinsi yang ada di Indonesia. Data yang didapatkan dari penelitian ini memiliki fitur yang sangat banyak dan bervariasi. Data tersebut juga memiliki dimensi yang jamak sehingga dapat digolongkan sebagai data *multidimensional*.

Dengan fakta diatas dapat diasumsikan, bila data tersebut akan sangat sulit dianalisa dengan hasil *output* yang sempurna. Dalam beberapa kondisi, data dengan dimensi tinggi dapat diasumsikan sebagai fungsi sampel yang mengharuskan proses *drill down* dan *roll up* pada tiap-tiap dimensi dari data tersebut. Pemilihan fitur merupakan bagian yang sangat penting dalam klasifikasi otomatis, dari banyak fitur yang ada pada data RISKESDAS 2007, maka akan digali kaidah yang ada di dalamnya untuk mendapatkan hasil analisa yang lebih tepat.

Riset Kesehatan Dasar (RISKESDAS) 2007 merupakan bagian dari Riset Kesehatan Dasar tingkat nasional yang dilakukan oleh Badan Penelitian dan Pengembangan Departemen Kesehatan RI dengan melibatkan BPS, organisasi profesi, perguruan tinggi, lembaga penelitian, pemerintah daerah, dan partisipasi masyarakat, untuk menyediakan informasi kesehatan yang berbasis bukti (*evidence-based*) untuk menunjang perencanaan bidang kesehatan kabupaten/Kota. Riskesdas mencakup sampel yang jauh lebih besar dari survei-survei kesehatan sebelumnya seperti SKRT atau SDKI

dan mencakup aspek kesehatan yang lebih luas. RISKESDAS 2007 dilaksanakan untuk menjawab pertanyaan tentang status kesehatan masyarakat di tingkat provinsi dan kabupaten/Kota, faktor-faktor yang melatarbelakanginya dan masalah kesehatan masyarakat yang spesifik di setiap wilayah[1].

Data mining adalah kegiatan yang meliputi pengumpulan, pemakaian data historis untuk menemukan keteraturan, pola atau hubungan dalam set data berukuran besar. Salah satu tugas dalam data mining adalah *Clustering*. Tujuan utama dari *Clustering* adalah pengelompokan sejumlah data/obyek ke dalam *cluster* sehingga dalam setiap *cluster* akan berisi data yang semirip mungkin. Dengan teknik ini, maka data tingkat kesehatan masyarakat tersebut dapat diklasterkan berdasarkan tingkatnya dan diharapkan dapat diperoleh informasi tersembunyi dari data set tersebut[2].

Metode *clustering* yang umum digunakan adalah K-Means Clustering, yang termasuk metode *partition clustering*, yakni memilah-milah data/obyek kedalam *cluster-cluster* yang ada. Menurut Jain, Metode k-means telah mengalami banyak pengembangan, antara lain: 1) Fuzzy-c-Means, 2) X-means, 3) k-Medoid, 4) Kernel k-means. Menurut Agusta, terdapat pula metode K-Harmonic MEans dan K-Modes. Variasi metode K-Means tersebut umumnya berhubungan dengan tiga hal yang telah disebutkan oleh Agusta. Namun saat ini telah dihasilkan pengembangan k-means clustering berbasis OWA oleh Cheng DKK yang melakukan *clustering* nilai agregat, yang merupakan kumpulan dari nilai multi atribut yang ada. Metode ini berbeda dengan variasi metode k-means yang telah ada sebelumnya. Pada penelitian cheng DKK lebih fokus pada cara untuk mengurangi kompleksitas data set eksperimental dan keterkaitan antara berbagai kriteria yang ada, yang dapat diatasi dengan cara menggabungkan k-means dengan OWA[3].

Ditinjau dari sisi perkembangan metode *Ordered Weighted Averaging* (OWA), ada beberapa penelitian terdahulu yang telah menerapkan metode OWA ini pada kasus *clustering* varietas padi menggunakan modifikasi metode K-Means berbasis OWA. Yager yang pertama kali memperkenalkan tentang OWA, menyatakan bahwa OWA dapat diaplikasikan untuk menyelesaikan berbagai problem, OWA dapat diterapkan untuk memberikan bobot yang berbeda untuk setiap atribut. OWA operator yang tepat dapat

digunakan untuk mewakili hubungan antar kriteria yang diintegrasikan. Pernyataan ini didukung oleh penelitian millatul ulya, yang telah Mengintegrasikan OWA dan K-Means *clustering* untuk mengelompokkan jenis-jenis padi berdasarkan preferensi konsumen, yaitu kadar amilosa, suhu, glatinisasi, persen beras kepala, persen beras putih, panjang bulir beras, bentuk bulir beras, pengapuran dan kadar protein[4].

OWA yang digunakan dalam penelitian Cheng DKK ini menggunakan persamaan OWA yang dikembangkan oleh Fuller and Majlender, dimana persamaan-persamaan tersebut tidak memperhatikan apakah data yang kita analisis memiliki label (output/variabel respon) atau tidak. karena persamaan ini hanya memerlukan dua input parameter saja yaitu jumlah variabel dari data dan nilai *orness* (α) atau parameter situasi yang digunakan. Oleh karena itu, K-means berbasis OWA ini digunakan dalam menyelesaikan kasus *clustering* data, seperti pada kasus data tingkat kesehatan masyarakat. Problem pada data RISKESDAS 2007 adalah memiliki variable yang multi atribut. OWA operator dapat mengurangi kompleksitas data dengan memadukan nilai-nilai multi atribut ke nilai-nilai agregat yang berupa nilai tunggal[5].

Kasus ini mirip dengan kasus pengambilan keputusan multikriteria (*Multicriteria Decision Making/MCDM*), dimana keputusan tingkat kesehatan masyarakat dari beberapa atribut didasarkan pada multikriteria. *Ordered Weighted Averaging* (OWA) sangat berguna untuk proses MCDM yang sering kali memerlukan keterkaitan antar kriteria yang ada[6]. Oleh karena itu, dengan menggunakan metode k-means berbasis OWA diharapkan hasil pengelompokan data RISKESDAS 2007 pada penelitian ini akan lebih optimal dibandingkan dengan hanya menggunakan metode K-Means *Clustering*.

1.2 Rumusan Masalah

Dari latar belakang yang telah dijelaskan sebelumnya dapat didapatkan suatu permasalahan sebagai berikut:

1. Bagaimana menerapkan metode K-Means berbasis OWA pada data RISKESDAS 2007 dengan mengukur tingkat performansinya menggunakan *Sum of Squares Error* (SSE)?

2. Membandingkan Nilai *Sum of Squares Error* antara metode K-Means berbasis OWA dengan K-Means *Clustering* pada data RISKESDAS 2007?

1.3 Tujuan Penelitian

Sesuai dengan rumusan masalah di atas, maka penelitian ini bertujuan sebagai berikut:

1. Mengaplikasikan K-Means berbasis OWA pada data RISKESDAS 2007 dan mengukur tingkat akurasi metode k-means OWA dalam menentukan tingkat kesehatan masyarakat menggunakan SSE (*Sum of Squares Error*).
2. Membandingkan Nilai *Sum of Squares Error* antara metode K-Means berbasis OWA dengan K-means pada data RISKESDAS 2007.

1.4 Manfaat Penelitian

Setelah penulis menyelesaikan penelitian yang berjudul pengelompokan tingkat kesehatan masyarakat menggunakan metode K-Means berbasis OWA, maka diharapkan nantinya dapat berguna dan bermanfaat. Manfaat penelitian ini diantaranya sebagai berikut:

1. Penelitian ini digunakan sebagai materi penunjang untuk penelitian selanjutnya yang dilakukan oleh Badan Penelitian dan Pengembangan Kesehatan (LITBANGKES) di Surabaya, yaitu penelitian data RISKESDAS 2013.
2. Penelitian ini digunakan sebagai referensi bagi mahasiswa Universitas Trunojoyo dan mahasiswa universitas lainya untuk mengembangkan penelitian ini atau membuat penelitian yang sesuai dengan bidang minat penelitian ini.
3. Penelitian yang telah dibuat ini, guna memperoleh gelar Strata-1.

1.5 Batasan Masalah

Untuk membatasi ruang lingkup yang akan dibahas pada penelitian ini, kami telah membuat batasan masalah sebagai berikut:

1. Data yang digunakan adalah data yang sudah di olah dari Badan Penelitian dan Pengembangan Kesehatan (LITBANGKES) di

Surabaya dengan menilai tingkat kesehatan masyarakat tiap Kabupaten se-Jawa Timur tahun 2007.

2. Metode yang digunakan pada penelitian ini adalah metode *Clustering* yaitu metode K-Means *Clustering* Berbasis OWA.

1.6 Metode Penelitian

Penelitian dan pengerjaan Tugas Akhir ini terbagi dalam beberapa tahap sebagai berikut :

1. Studi Pustaka

Proses pengumpulan data dan informasi dengan mempelajari materi-materi tentang *clustering* menggunakan metode K-Means Clustering dan OWA (*Ordered Weighted Averaging*) dan kajian-kajian lain yang terkait sebagai pendukung.

2. Survei

Pada tahap ini akan dilakukan survei tentang bagaimana proses pengolahan data yang dilakukan oleh instansi terkait saat mengolah data RISKESDAS 2007.

3. Perumusan Masalah

Pada tahap ini yang dikehjkan adalah merumuskan masalah yang didapat, dan menentukan kebutuhan sistem yang tepat untuk mengatasinya.

4. Analisa dan Perancangan

Pada tahap ini proses analisa system dan mulai dilakukan perancangan design dari system yang akan digunakan.

5. Implementasi

Pada proses ini dilakukan pengimplementasian metode K-Means *Clustering* dan OWA (*Ordered Weighted Averaging*) untuk menganalisa data yang sudah didapat.

6. Evaluasi

Pada tahapan ini adalah melakukan evaluasi terhadap sistem yang dibuat. Apabila terjadi kekurangan maka akan dilakukan perbaikan.

7. Penulisan Laporan

Tahap ini merupakan tahap terakhir dari penyusunan Tugas Akhir. Pada tahap ini dilakukan penulisan laporan dari Tugas Akhir yang dikerjakan.

1.7 Sistematika Penulisan

Dalam laporan Tugas Akhir ini akan dibagi menjadi lima bagian pembahasan, dari lima pembahasan tersebut di uraikan sebagai berikut:

BAB I PENDAHULUAN

Pada bagian ini dipaparkan mengenai hal-hal yang melatar belakangi penelitian, rumusan permasalahan, dan batasan permasalahan dalam penelitian, serta tujuan dan relevansi atau manfaat penelitian terhadap perkembangan solusi dari permasalahan, serta sistematika penulisan dalam penulisan laporan Tugas Akhir ini.

BAB II TINJAUAN PUSTAKA

Dalam bab ini diuraikan teori-teori penunjang dalam penulisan laporan Tugas Akhir ini. Pustaka yang digunakan baik dari buku, diktat, maupun materi dari internet.

BAB III PERANCANGAN SISTEM

Dalam bagian ini diuraikan secara rinci rancangan sistem dari aplikasi Tugas Akhir ini.

BAB IV SIMULASI DAN ANALISIS

Pada bagian ini diuraikan secara rinci langkah-langkah yang dilakukan dalam penelitian dalam lingkup simulasi. Serta analisis sistem hasil simulasi tersebut.

BAB V PENUTUP

Pada bab ini dijelaskan hasil dari simulasi, yang termaktub di dalam kesimpulan. Dan juga saran untuk mengembangkan hasil analisis penelitian Tugas Akhir ini agar dapat dikembangkan ke penelitian berikutnya.