

BAB I

PENDAHULUAN

1.1 Latar Belakang Masalah

Jagung merupakan tanaman yang banyak dijumpai di daerah tropis. Jagung juga tergolong tanaman pangan terpenting dunia setelah gandum dan padi. Beberapa daerah di Indonesia (misalnya di Madura dan Nusa Tenggara) merupakan salah satu penghasil jagung sebagai pangan pokok. Ada beberapa jenis jagung yang tumbuh di Indonesia dan termasuk jenis jagung yang dihasilkan dari rekayasa genetika (Hartono, 2006).

Mengingat meningkatnya permintaan pasar *domestic* ataupun internasional yang sangat besar terhadap kebutuhan jagung untuk bahan pakan ternak rata-rata 950.000 ton per bulan, terdiri dari 700.000 untuk industri pakan dan 250.000 untuk peternak mandiri. Adapun, produksi jagung meningkat signifikan sejak 2015 sebesar 19.6 juta ton, menjadi 23.5 juta ton pada 2016 dan 27.9 juta ton pada 2017. Menteri Pertanian Amran Sulaiman mengatakan pemerintah akan menggarap peluang ekspor jagung ke Malaysia sekitar 3 juta ton - 4 juta ton dan Filipina 1 juta ton di tahun 2018 (Bisnis.com, 2018). Adanya permintaan jagung yang sangat besar ini memicu para peneliti untuk menghasilkan bakal jagung yang berkualitas tinggi dengan menyaring atau mengelompokkan varietas-varietas jagung unggulan yang diperoleh dari informasi masing-masing gene jagung.

Melalui ekspresi gen, peneliti dapat menganalisis varietas-varietas jagung dan mengelompokkannya berdasarkan informasi genetik, dari penelitian ini banyak menghasilkan data biologis yang harus dihitung dan diinterpretasi secara akurat. Bioinformatika merupakan ilmu yang mengkombinasikan ilmu komputer, statistik, matematika, teknik analisis

dan menginterpretasi data biologi. Bioinformatika telah diaplikasikan pada berbagai penelitian, termasuk dalam studi genetika tanaman jagung.

Penelitian yang dilakukan oleh Andayani *et al.* (2017) dalam penelitiannya bertujuan untuk mempelajari bioinformatika genetika jagung kaya vitamin A dan melakukan pengelompokan genetik inbrida jagung provit-A. Sementara Liu *et al.* (2005) melalui studi anotasi gen yang diperbaiki, karakterisasi gen-gen, dan serangkaian inti gen kandidat untuk lebih lanjut mengkarakterisasi perkembangan reproduksi jagung dan meningkatkan potensi hasil biji-bijian. Kemudian penelitian yang dilakukan oleh Lawrence *et al.* (2005) mengembangkan *database* genom jagung yang menyediakan informasi genetik dengan *interface* berbasis *website*.

Selanjutnya Penelitian yang dilakukan oleh Jakt *et al.* (2001) melakukan penelitian ekspresi gen berskala besar dan proyek sekuensing genom memberikan banyak informasi yang dapat digunakan untuk mengidentifikasi atau memprediksi proses pengaturan seluler. Gen dapat dikelompokkan berdasarkan kesamaan profil atau fungsi ekspresi dan kelompok ini cenderung mengandung gen yang diatur oleh faktor transkripsi yang sama.

Mengingat analisis bioinformatika dalam studi genetik tanaman jagung, memperoleh banyak data biologis yang berskala besar dalam bentuk matriks, maka peneliti membutuhkan pendekatan komputasi yaitu *Data mining*. *Data mining* adalah proses yang menggunakan teknik statistik, matematika, kecerdasan buatan, dan *machine learning* untuk mengekstraksi dan mengidentifikasi informasi yang bermanfaat dan pengetahuan yang terkait dari berbagai *database* besar. Salah satu teknik yang ada pada *data mining* adalah *Biclustering*. Teknik *Biclustering* merupakan teknik baru yang berkembang dari teknik *cluster* pada umumnya (Kusrini, 2009).

Biclustering merupakan teknik pengelompokan data, dimana baris dikelompokkan dan dalam waktu yang sama juga dilakukan

pengelompokan berdasarkan variabel. Pada data terstruktur ini berarti baris dan kolom dikelompokkan secara bersamaan. *Biclustering* adalah metode analisis data yang dirancang untuk mendeteksi pola lokal dalam data. Metode *biclustering* terdapat banyak algoritma yang telah digunakan untuk pengelompokan dan mengidentifikasi pola dalam data biologis yaitu Cheng dan church, Plaid, OPSM, ISA, Spectral, xMotifs dan Bimax.

Adapun penelitian sebelumnya yang menjelaskan tentang metode Algoritma *biclustering* xMotif yang diusulkan oleh Murali dan Kasif (2003), *Bicluster* dengan evolusi yang koheren diwakili oleh algoritma xMotifs. Algoritma ini mencari baris dengan nilai konstan di atas satu set kolom. Untuk data ekspresi gen, mereka menyebut *bicluster* "motif ekspresi gen yang diawetkan", disingkat menjadi "xMotif". Untuk aplikasi ini sangat penting untuk menemukan metode *preprocessing* yang baik dikarenakan tujuan utama dari algoritma ini adalah untuk mendefinisikan status gen (baris) yang sama dalam sampel yang dipilih (kolom), ini disebut gen yang dilestarikan (baris). Salah satu cara untuk menangani keadaan gen adalah dengan hanya mendiskritkan data.

Pada penelitian ini peneliti tertarik untuk menganalisis ekspresi gen jagung berskala besar dan gen-gen dapat di kelompokkan berdasarkan kesamaan profil atau fungsi ekspresi yaitu, dengan judul “APLIKASI ALGORITMA *MOTIF* EKSPRESI (xMotif) *BICLUSTERING* PADA DATA *GENE EXPRESSION MAIZE GROWTH STAGES*”

1.2 Rumusan Masalah

1. Bagaimana melakukan analisis *biclustering* pada data ekspresi gen pertumbuhan jagung dengan algoritma xMotif?
2. Bagaimana melakukan evaluasi terhadap hasil *biclustering* dengan algoritma xMotif?
3. Bagaimana implikasi *biclustering* dengan algoritma xMotif pada data ekspresi gen pertumbuhan jagung?

1.3 Batasan Masalah

Adapun batasan masalah dalam penelitian ini adalah :

1. Bahasa pemrograman dan *software* yang digunakan dalam penelitian ini adalah *software* R 3.4.3 dengan beberapa *package bioinformatics*.
2. Data yang digunakan dalam penelitian ini adalah data *maize gene expression* dengan kode akses PRJNA343268.
3. Algoritma yang digunakan xMotif (*expression motif*).

1.4 Jenis Penelitian dan Metode Analisis

Jenis penelitian yang dilakukan adalah aplikasi, menggunakan analisis *Biclustering* dengan algoritma xMotif pada data ekspresi gen pertumbuhan jagung.

1.5 Tujuan Penelitian

1. Mengetahui analisis *biclustering* pada data ekspresi gen pertumbuhan jagung dengan algoritma xMotif.
2. Mengetahui evaluasi terhadap hasil *biclustering* dengan algoritma xMotif.
3. Mengetahui implikasi *biclustering* dengan algoritma xMotif pada data ekspresi gen pertumbuhan jagung.

1.6 Manfaat Penelitian

Pada Penulisan skripsi ini diharapkan dapat memberikan manfaat sebagai berikut:

1. Bagi Penulis
Dapat menambah pengetahuan tentang analisis *biclustering* khususnya dengan metode xMotif.
2. Bagi Pembaca
Penelitian ini dapat dijadikan sebagai referensi pembaca untuk mengembangkan metode xMotif pada kasus-kasus yang lainnya.
3. Bagi Instansi

Didapatkan gambaran hasil *biclustering* varietas jagung pada data ekspresi gen tanaman jagung diharapkan dapat menjadi bahan pertimbangan baik Instansi Pemerintah seperti dinas pertanian untuk pemuliaan jagung.