

HALAMAN PERSETUJUAN PEMBIMBING

TUGAS AKHIR

Judul : Implementasi *Biclustering* Dengan Algoritma *Plaid* Pada
Data *Maize Growth Stage*

Nama Mahasiswa : Shodiq Muhammad

Nomor Mahasiswa: 14 611 086

**TUGAS AKHIR INI TELAH DIPERIKSA DAN DISETUJUI UNTUK
DIUJIKAN**

Yogyakarta, 29 Juli 2018

Pembimbing I,



(Dr. techn. Rohmatul Fajriyah, S.Si., M.Si)

Pembimbing II,



(Husna Nugrahapraja, Ph.D)

HALAMAN PENGESAHAN
TUGAS AKHIR
IMPLEMENTASI BICLUSTERING DENGAN ALGORITMA PLAID
PADA DATA MAIZE GROWTH STAGE

Nama Mahasiswa : Shodiq Muhammad
Nomor Mahasiswa : 14 611 086

TUGAS AKHIR INI TELAH DIUJIKAN
PADA TANGGAL 31 JULI 2018

Nama Penguji

1. Muryanto, S.P., M.Si
2. Ayundyah Kesunawati, S.Si., M.Si
3. Husna Nugrahapraja, Ph.D
4. Dr. techn. Rohmatul Fajriyah, S.Si., M.Si

Tanda Tangan



Disetujui,
Dekan Fakultas Matematika dan Ilmu Pengetahuan Alam



(Shodiq Muhammad, S.Si., M.Si., Ph.D)

KATA PENGANTAR



Assalamu'alaikum Wr. Wb.

Alhamdulillah, segala puji dan syukur penulis haturkan kepada Allah SWT yang telah memberikan nikmat yang tak terhitung sehingga penulis dapat menyelesaikan tugas akhir yang berjudul **“Implementasi Biclustering Dengan Algoritma Plaid Pada Data Maize Growth Stage”**. Shalawat serta salam penulis haturkan kepada teladan terbaik Nabi Muhammad SAW beserta keluarga dan sahabat-sahabatnya, semoga penulis dan kita semua mendapatkan pertolongan di akhirat nanti.

Tugas akhir ini disusun untuk memenuhi salah satu syarat memperoleh gelar Sarjana Statistika di Program Studi Statistika Fakultas Matematika dan Ilmu Pengetahuan Alam Universitas Islam Indonesia. Penulis menyadari bahwa tanpa bantuan dan motivasi dari berbagai pihak, penulisan tugas akhir ini tidak dapat diselesaikan dengan baik. Terimakasih penulis sampaikan kepada:

1. Kedua orang tua tercinta yang selalu mendoakan penulis dan memberikan amanah serta kepercayaan ini. Kedua adik saya, Mahmudi dan Latifah, serta saudara-saudari tercinta yang selalu mendukung, memberi semangat, mendorong, dan mendoakan yang terbaik bagi penulis.
2. Bapak Prof. Riyanto, S.Pd., M.Si., Ph.D., selaku Dekan Fakultas Matematika dan Ilmu Pengetahuan Alam Universitas Islam Indonesia.
3. Bapak RB Fajriya Hakim, M.Si., selaku Ketua Jurusan Statistika beserta seluruh jajarannya.
4. Ibu Dr. techn. Rohmatul Fajriyah, S.Si., M.Si selaku pembimbing pertama yang telah membimbing dan senantiasa mendoakan penulis. Semoga Allah memberikan kesehatan kepada ibu dan ilmu yang diberikan bermanfaat bagi penulis.

5. Bapak Husna Nugrahapraja, Ph.D., selaku pembimbing kedua, terimakasih telah banyak memberikan wawasan ilmu pengetahuan baru di bidang bioinformatika. Semoga segala keikhlasannya dibalas oleh Allah sebagai amal jariyah.
6. Ibu Atina, Pak Hasan, Pak Muhajir, serta Bapak dan Ibu Dosen Program Studi Statistika lainnya, terimakasih ilmu yang diberikan semoga senantiasa diberikan kebaikan oleh Allah.
7. Keluarga besar Statistka FMIPA UII, khususnya teman-teman Statistika angkatan 2014 yang telah memberikan banyak kenangan. Sukses untuk kita semua.
8. Teman seperjuangan satu bimbingan, Mas Rakhmat, Husain, Aat, Ari, Gilang, Hawila, Nanda, Lina, Ummi, Ika, Leni, Himelda, Maidah semoga kita dapat selalu menebar manfaat dimanapun.
9. Anggota grup *ngaji apa ngaji* yang tidak pernah lelah saling berbagi nasihat. Semoga kelak kita dipertemukan kembali dalam keadaan yang jauh lebih baik. Tetap istiqomah.
10. Teman-teman yang selalu bertanya “sidang kapan?” semoga dengan ini kalian mendapatkan jawaban apa yang kalian tanyakan.
11. Semua pihak yang tidak bisa disebutkan satu persatu, terimakasih atas segala bantuannya.

Penulis menyadari bahwa tugas akhir ini masih jauh dari sempurna. Oleh karena itu, kritik dan saran dari semua pihak yang bersifat membangun selalu penulis harapkan demi kesempurnaan Tugas Akhir ini. Semoga Tugas Akhir ini dapat bermanfaat. Akhir kata, semoga Allah SWT senantiasa meridhai segala usaha kita. *Aamiin yaa rabbal'alamiin.*

Wassalamu'alaikum, Wr. Wb.

Yogyakarta, Agustus 2018

Penulis

DAFTAR ISI

HALAMAN PERSETUJUAN PEMBIMBING	Error! Bookmark not defined.
HALAMAN PENGESAHAN	Error! Bookmark not defined.
KATA PENGANTAR	iii
DAFTAR ISI	v
DAFTAR TABEL	vii
DAFTAR GAMBAR	viii
DAFTAR LAMPIRAN	ix
SURAT PERNYATAAN	Error! Bookmark not defined.
INTISARI	x
ABSTRACT	xii
BAB I PENDAHULUAN	Error! Bookmark not defined.
1.1. Latar Belakang	Error! Bookmark not defined.
1.2. Rumusan Masalah	Error! Bookmark not defined.
1.3. Batasan Masalah	Error! Bookmark not defined.
1.4. Tujuan Penelitian	Error! Bookmark not defined.
1.5. Manfaat Penelitian	Error! Bookmark not defined.
BAB II TINJAUAN PUSTAKA	Error! Bookmark not defined.
3.1. Penelitian Mengenai Metode <i>Plaid</i>	Error! Bookmark not defined.
3.2. Penelitian <i>Biclustering</i> Pada Jagung	Error! Bookmark not defined.
3.3. Penelitian Metode <i>Biclustering</i> Menggunakan R	Error! Bookmark not defined.
BAB III LANDASAN TEORI	Error! Bookmark not defined.
3.1. Botani Tanaman Jagung	Error! Bookmark not defined.
3.1.1. Taksonomi	Error! Bookmark not defined.
3.1.2. Morfologi	Error! Bookmark not defined.
3.2. Pemuliaan Tanaman	Error! Bookmark not defined.
3.3. Bioinformatika	Error! Bookmark not defined.
3.6.1 Pengertian Bioinformatika	Error! Bookmark not defined.

3.6.2	<i>Next Generation Sequencing (NGS)</i>	Error! Bookmark not defined.
3.6.3	<i>Gene Expression Analysis</i>	Error! Bookmark not defined.
3.4.	<i>Biclustering</i>	Error! Bookmark not defined.
3.4.1	Pengertian <i>Biclustering</i>	Error! Bookmark not defined.
3.4.2	Tipe dan struktur <i>biclustering</i>	Error! Bookmark not defined.
3.5.	<i>Filtering</i>	Error! Bookmark not defined.
3.6.	<i>Feature Selection</i>	Error! Bookmark not defined.
3.7.	Model <i>Plaid</i>	Error! Bookmark not defined.
3.8.	<i>Software R</i>	Error! Bookmark not defined.
BAB IV METODOLOGI PENELITIAN		Error! Bookmark not defined.
4.1.	Tempat dan Waktu Penelitian	Error! Bookmark not defined.
4.2.	Sumber Data	Error! Bookmark not defined.
4.3.	Metode Analisis Data	Error! Bookmark not defined.
4.4.	Variabel dan Definisi Operasional Variabel	Error! Bookmark not defined.
4.5.	Langkah Penelitian	Error! Bookmark not defined.
BAB V HASIL DAN PEMBAHASAN		Error! Bookmark not defined.
5.1.	Pendahuluan	Error! Bookmark not defined.
5.2.	Ringkasan dan Struktur Data	Error! Bookmark not defined.
5.3.	Analisis Deskriptif	Error! Bookmark not defined.
5.4.	<i>Filtering</i>	Error! Bookmark not defined.
5.5.	Algoritma <i>Plaid</i>	Error! Bookmark not defined.
BAB VI KESIMPULAN DAN SARAN		Error! Bookmark not defined.
LAMPIRAN		Error! Bookmark not defined.

DAFTAR TABEL

Tabel 2.1 Pembagian kelas <i>biclustering</i>	7
Tabel 3.1 Ringkasan perbedaan NGS dan Sanger <i>sequencing</i>	17
Tabel 3.2 Matriks gen ekspresi.....	19
Tabel 4.1 Definisi Variabel Operasional.....	28
Tabel 5.1 Daftar <i>packages</i> yang diperlukan pada proses <i>install</i>	31
Tabel 5.2 Ringkasan informasi data.....	34
Tabel 5.3 Hasil <i>bicluster</i> yang terbentuk.....	36
Tabel 5.4 Karakteristik gen pada tiap tiap <i>bicluster</i>	41

DAFTAR GAMBAR

Gambar 2.1 Struktur <i>BiclustGUI</i> Packages.....	11
Gambar 3.1 Konsep perbedaan antara hasil analisis <i>cluster</i> (kiri) dan <i>bicluster</i> (kanan).....	20
Gambar 3.2 Contoh tipe tipe <i>bicluster</i> . (a) nilai konstan, (b) baris konstan, (c) kolom konstan, (d) nilai koheren (model aditif), (e) nilai koheren (model multiplikatif).....	21
Gambar 3.3 Struktur <i>bicluster</i> . (a) <i>single bicluster</i> , (b) <i>exclusive</i> baris dan kolom <i>bicluster</i> , (c) struktur <i>checkboxboard</i> , (d) <i>exclusive</i> baris <i>bicluster</i> , (e) <i>exclusive</i> kolom <i>bicluster</i> , (f) <i>non-overlapping bicluster</i> dengan struktur <i>tree</i> , (g) <i>non-overlapping non-exclusive bicluster</i> , (h) <i>overlapping bicluster</i> dengan <i>hierarchical</i> , (i) <i>arbitrarily positioned overlapping bicluster</i>	21
Gambar 4.1 Diagram alur penelitian.....	29
Gambar 5. 1 Banyaknya gen tiap kromosom.....	34
Gambar 5. 2 Barchart rata-rata ekspresi gen setiap variabel pada organ.....	35
Gambar 5. 3 <i>Heatmap bicluster 1</i>	37
Gambar 5. 4 <i>Heatmap bicluster 2</i>	37
Gambar 5. 5 <i>Heatmap bicluster 3</i>	38
Gambar 5. 6 <i>Heatmap bicluster 4</i>	39
Gambar 5.7 <i>Heatmap bicluster 5</i>	39
Gambar 5.8(a) Plot rata-rata <i>bicluster 1</i>	40
Gambar 5.8(b) Plot rata-rata <i>bicluster 2</i>	40
Gambar 5.8(c) Plot rata-rata <i>bicluster 3</i>	40
Gambar 5.8(d) Plot rata-rata <i>bicluster 4</i>	40
Gambar 5.8(e) Plot rata-rata <i>bicluster 5</i>	40

DAFTAR LAMPIRAN

Lampiran 1. Script R Filtering	61
Lampiran 2. Algoritma Plaid	63
Lampiran 3. Hasil <i>biclustering</i> algoritma plaid	64

SURAT PERNYATAAN

Dengan ini saya menyatakan bahwa dalam Tugas Akhir ini tidak terdapat karya yang sebelumnya pernah diajukan untuk memperoleh gelar kesarjanaan di suatu perguruan tinggi dan sepanjang pengetahuan saya juga tidak terdapat karya atau pendapat yang pernah ditulis atau diterbitkan oleh orang lain, kecuali yang diacu dalam naskah ini dan disebutkan dalam daftar pustaka.

Yogyakarta, Juli 2018



IMPLEMENTASI *BICLUSTERING* DENGAN ALGORITMA *PLAID* PADA DATA *MAIZE GROWTH STAGE*

Oleh: Shodiq Muhammad

Program Studi Statistika Fakultas Matematika dan Ilmu Pengetahuan Alam

Universitas Islam Indonesia

E-mail: 14611086@students.uii.ac.id

INTISARI

Data ekspresi gen merupakan teknologi yang handal dalam menjelaskan tentang fungsi genomik dalam ilmu kehidupan. Pada kasus ini, kami menggunakan data ekspresi gen maize growth stage yang didapat dari Maize qTeller. Data tersebut di analisis dengan biclustering algoritma plaid untuk mencari clustering secara bersamaan pada dimensi baris dan kolom. Sebelum di analisis algoritma plaid, terlebih dahulu dilakukan filtering untuk menghilangkan missing value dan uji t untuk mengurangi baris yang memiliki p-value dibawah 0.0001. Analisis algoritma plaid menggunakan paket RcmdrPlugin.BiclustGUI di R. Kami menggunakan beberapa parameter yaitu shuffle=3, back.fit=0, max.layers=20, iter.startup=5, iter.layer=10. Maka dapat diketahui hasil 5 bicluster di antaranya bicluster 1 berdimensi 2 gen dan 35 organ, bicluster 2 berdimensi 2 gen dan 16 organ, bicluster 3 berdimensi 6 gen dan 30 organ, bicluster 4 berdimensi 3 gen dan 10 organ, bicluster 5 berdimensi 47 gen dan 24 organ.

Kata Kunci: *biclustering*, tanaman jagung, algoritma *plaid*

IMPLEMENTATION *BICLUSTERING* WITH *PLAID* ALGORITHM IN *MAIZE GROWTH STAGE DATASET*

Shodiq Muhammad

Departement of Statistics, Faculty of Mathematics and Natural Sciences Islamic
Universit of Indonesia

E-mail: 14611086@students.uii.ac.id

ABSTRACT

Gene expression data is a reliable technology in explaining the function of genomics in life sciences. In this case, we used the expression data on the growth stage of the corn gene obtained from Corn qeller. The data is processed by biclustering plaid algorithm to find groupings simultaneously in dimensions and columns. Before doing the plaid algorithm, filtering is done first to remove the missing value and t test to reduce the line that has a p-value below 0.0001. Plaid algorithm analysis uses the RcmdrPlugin package. BiclustGUI in R. We use several parameters namely shuffle = 3, back.fit = 0, max.layers = 20, iter.startup = 5, iter.layer = 10. Then 5 bicluster results can be informed Bicluster 1 was explored with 2 genes and 35 organs, bicluster 2 with 2 genes and 16 organs, bicluster 3 with 6 genes and 30 organs, bicluster 4 with 3 genes and 10 organs, bicluster 5 with 47 genes and 24 organs.

Kata Kunci: *biclustering, Maize, algorithm plaid*

