

**PERBANDINGAN NILAI AKURASI KLASIFIKASI KANKER
PAYUDARA DENGAN DATA *MICRORNA* MENGGUNAKAN
METODE *DEEP NEURAL NETWORK* (DNN) TANPA DAN DENGAN
OPTIMASI *BATCH NORMALIZATION***



SKRIPSI

**Disusun Sebagai Salah Satu Syarat
untuk Memperoleh Gelar Sarjana Komputer
pada Departemen Ilmu Komputer/ Informatika**

Disusun oleh :

ANGGUN CAHYA NINGRUM

24010314120007

**DEPARTEMEN ILMU KOMPUTER/INFORMATIKA
FAKULTAS SAINS DAN MATEMATIKA
UNIVERSITAS DIPONEGORO**

2018

HALAMAN PERNYATAAN KEASLIAN SKRIPSI

Saya yang bertanda tangan di bawah ini:

Nama : Anggun Cahya Ningrum
NIM : 24010314120007
Judul : Perbandingan Nilai Akurasi Klasifikasi Kanker Payudara dengan Data *MicroRNA* Menggunakan Metode *Deep Neural Network* (DNN) Tanpa dan Dengan Optimasi *Batch Normalization*.

Dengan ini saya menyatakan bahwa dalam tugas akhir/skripsi ini tidak terdapat karya yang 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 tertulis diacu dalam naskah ini dan disebutkan di dalam daftar pustaka.

Semarang, 3 Januari 2019



Anggun Cahya Ningrum
24010314120007

HALAMAN PENGESAHAN

Judul : Perbandingan Nilai Akurasi Klasifikasi Kanker Payudara dengan Data
MicroRNA Menggunakan Metode *Deep Neural Network* (DNN) Tanpa
dan Dengan Optimasi *Batch Normalization*
Nama : Anggun Cahya Ningrum
NIM : 24010314120007

Telah diujikan pada sidang skripsi pada tanggal 13 Desember 2018 dan dinyatakan lulus
pada tanggal 13 Desember 2018

Mengetahui,
Ketua Departemen Ilmu Komputer/ Informatika



Semarang, 3 Januari 2019

Panitia Penguji Tugas Akhir
Ketua,

Drs. Eko Adi Sarwoko, M.Kom
NIP. 196511071992031003

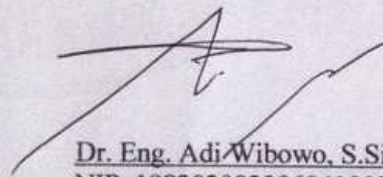
HALAMAN PENGESAHAN

Judul : Perbandingan Nilai Akurasi Perbandingan Nilai Akurasi Klasifikasi Kanker Payudara dengan Data *MicroRNA* Menggunakan Metode *Deep Neural Network* (DNN) Tanpa dan Dengan Optimasi *Batch Normalization*
Nama : Anggun Cahya Ningrum
NIM : 24010314120007

Telah diujikan pada sidang skripsi pada tanggal 13 Desember 2018.

Semarang, 3 Januari 2019

Pembimbing,



Dr. Eng. Adi Wibowo, S.Si., M.Kom
NIP. 198203092006041002

ABSTRAK

Kanker payudara merupakan salah satu penyebab kematian terbanyak di dunia. Salah satu upaya yang dapat dilakukan untuk mengurangi jumlah kasus kematian akibat kanker ini adalah melakukan deteksi dini dengan menggunakan data *MicroRNA*. *MicroRNA* merupakan salah satu *biomarker* kanker yang dapat membantu dalam proses klasifikasi. *MicroRNA* dapat digunakan untuk mengenali apakah suatu sel merupakan sel kanker atau bukan bahkan dalam stadium paling awal. Metode *Deep Neural Network* (DNN) terdiri dari dua atau lebih lapisan belajar mandiri (*hidden layer*), dengan bobot unit yang terhubung sepenuhnya antara dua lapisan yang berdekatan dapat dipelajari secara otomatis. Namun, DNN masih memiliki kelemahan yaitu perubahan dalam distribusi *input* setiap lapisan yang menyebabkan masalah karena lapisan perlu terus beradaptasi dengan distribusi baru sehingga menghasilkan nilai akurasi yang kurang optimal. Dikarenakan perubahan distribusi *input* setiap lapisan ini, salah satu solusi untuk mengoptimalkan kinerja DNN adalah dengan menggunakan algoritma *Batch Normalization*. *Batch Normalization* menjadikan normalisasi sebagai bagian dari arsitektur model yang membuat jaringan lebih stabil dengan rentang gradient yang lebih kecil. Penelitian ini dilakukan untuk mengoptimalkan kinerja metode DNN yang digunakan untuk klasifikasi kanker payudara menggunakan data *MicroRNA* dengan menambahkan algoritma *Batch Normalization*. Hasil penelitian pada DNN standar dengan 4 *hidden layer* dengan masing-masing 200 *hidden unit*, dan *dropout rate* 50% menghasilkan rata-rata nilai akurasi sebesar 58.67%. Sedangkan hasil penelitian menunjukkan bahwa penerapan algoritma *Batch Normalization* pada DNN standar dengan 2 *hidden layer* dengan masing-masing 200 *hidden unit*, dan *dropout rate* 50% menghasilkan rata-rata nilai akurasi terbaik yaitu sebesar 93.67%. Menggunakan data *MicroRNA* kanker payudara, penerapan metode DNN dengan algoritma *Batch Normalization* terbukti mampu meningkatkan akurasi DNN standar untuk jumlah *hidden layer* kecil.

Kata Kunci : Kanker payudara, *MicroRNA*, *Deep Neural Network*, *Batch Normalization*

ABSTRACT

Breast cancer is one of the most common causes of death in the world. One way that can be done to reduce the number of death cases are to do early detection using MicroRNA data. MicroRNA is one of the cancer biomarkers that can help in the classification process. MicroRNA can be used to identify whether a cell is a cancer cell or not even in the earliest stages. Deep Neural Network (DNN) method consists of two or more layers of self-learning units (hidden units). The weight of hidden units that are fully connected between two layers can be learn automatically. However, DNN still has a weakness, namely changes in the distribution of each layer's inputs that cause problems, because the layers need to continue to adapt to the new distribution and produce less optimal accuracy values. Due to changes in the distribution of each layer's inputs, one solution to optimize DNN's performance is to use the Batch Normalization algorithm. Batch Normalization makes normalization a part of the model architecture that make the network more stable, where the gradient has a smaller range. This research was conducted to optimize the performance of the DNN method used for the classification of breast cancer by adding a batch normalization algorithm. The results showed that the implementation of the standard DNN with 4 hidden layers, 200 hidden units, and a 50% dropout rate resulted the average accuracy value of 58.67%. The implementation of Batch Normalization algorithm on a standard DNN with 2 hidden layers with each 200 hidden units, and a dropout rate of 50% produces the average value of the best accuracy of 93.67%. By using MicroRNA breast cancer data, the DNN method with the Batch Normalization algorithm is proven to be able to increase the accuracy of standard DNN for a small number of hidden layers.

Keywords : Breast Cancer, MicroRNA, Deep Neural Network, Batch Normalization

KATA PENGANTAR

Puji syukur penulis panjatkan kehadirat Tuhan Yang Maha Esa yang telah melimpahkan segala rahmat dan hidayah-Nya sehingga penulis dapat menyelesaikan skripsi dengan judul “Perbandingan Nilai Akurasi Klasifikasi Kanker Payudara dengan Data *MicroRNA* Menggunakan Metode *Deep Neural network* (DNN) Tanpa dan Dengan Optimasi *Batch Normalization*”.

Skripsi ini dibuat dengan tujuan sebagai salah satu syarat untuk memperoleh gelar sarjana komputer pada Departemen Ilmu Komputer/Informatika Fakultas Sains dan Matematika Universitas Diponegoro, Semarang.

Dalam pelaksanaan skripsi serta penyusunan dokumen skripsi ini, penulis menyadari banyak pihak yang membantu sehingga akhirnya dokumen ini dapat diselesaikan. Oleh karena itu, melalui kesempatan ini penulis ingin menyampaikan ucapan terima kasih yang sebesar-besarnya kepada :

1. Ibu Dr. Retno Kusumaningrum, S.Si, M.Kom selaku Kepala Departemen Ilmu Komputer/Informatika Fakultas Sains dan Matematika Universitas Diponegoro, Semarang.
2. Bapak Helmie Arif Wibawa, S.Si, M.Cs, selaku Koordinator Skripsi Departemen Ilmu Komputer/Informatika Fakultas Sains dan Matematika Universitas Diponegoro, Semarang.
3. Dr. Eng. Adi Wibowo, S.Si, M.Kom, selaku dosen pembimbing skripsi yang telah membantu dalam membimbing dan mengarahkan penulis hingga selesainya skripsi ini.
4. Semua pihak yang telah membantu kelancaran dalam penyusunan skripsi, yang tidak dapat penulis sebutkan satu persatu.

Penulis menyadari bahwa dokumen skripsi ini masih jauh dari sempurna. Oleh karena itu, saran dan kritik yang membangun sangat penulis harapkan. Akhir kata, semoga skripsi ini dapat bermanfaat bagi semua pihak.

Semarang, 3 Januari 2019

Penulis

HALAMAN PERNYATAAN PERSETUJUAN PUBLIKASI SKRIPSI UNTUK KEPENTINGAN AKADEMIS

Sebagai civitas akademik Universitas Diponegoro, saya yang bertanda tangan di bawah ini:

Nama : Anggun Cahya Ningrum

NIM : 24010314120007

Program Studi : Ilmu Komputer/ Informatika

Jenis Karya : Skripsi

Demi pengembangan ilmu pengetahuan, menyetujui untuk memberikan **Hak Bebas Royalti Noneksklusif (*Non-exclusive Royalty Free Right*)** kepada Universitas Diponegoro atas karya ilmiah saya yang berjudul:

Perbandingan Nilai Akurasi Klasifikasi Kanker Payudara dengan Data *MicroRNA* Menggunakan Metode *Deep Neural Network* (DNN) Tanpa dan Dengan Optimasi *Batch Normalization*

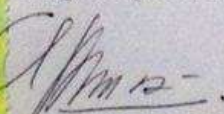
Beserta perangkat yang ada (jika diperlukan). Dengan Hak bebas Royalti Non-eksklusif ini Universitas Diponegoro berhak menyimpan, mengalihmedia/ formatkan, mengelola dalam bentuk pangkalan data (*database*), merawat, dan mempublikasikan skripsi saya tanpa meminta izin dari saya selama tetap mencantumkan nama saya sebagai penulis/ pencipta dan sebagai pemilik Hak Cipta.

Demikian pernyataan ini saya buat dengan sebenarnya

Semarang, 3 Januari 2019

Yang Menyatakan




Anggun Cahya Ningrum
2401031412007

DAFTAR ISI

HALAMAN PERNYATAAN KEASLIAN SKRIPSI.....	ii
HALAMAN PENGESAHAN	iii
HALAMAN PENGESAHAN	iv
ABSTRAK	v
ABSTRACT	vi
KATA PENGANTAR.....	vii
HALAMAN PERNYATAAN PERSETUJUAN PUBLIKASI SKRIPSI UNTUK KEPENTINGAN AKADEMIS.....	viii
DAFTAR ISI	ix
DAFTAR GAMBAR.....	xiii
DAFTAR TABEL	xv
BAB I PENDAHULUAN	1
1.1 Latar Belakang.....	1
1.2 Rumusan Masalah.....	3
1.3 Tujuan dan Manfaat.....	3
1.4 Ruang Lingkup	3
1.5 Sistematika Penulisan	4
BAB II LANDASAN TEORI.....	5
2.1. <i>MicroRNA</i> sebagai Potensi <i>Biomarker</i> Kanker	5
2.2. Klasifikasi	5
2.3. Metode <i>Deep Neural Network</i>	6
2.3.1 Arsitektur Metode <i>Deep Neural Network</i>	8
2.3.2 Fungsi Aktivasi <i>Rectified Linear Units</i>	8
2.3.3 Fungsi Aktivasi <i>Sigmoid</i>	9
2.3.4 <i>Batch Normalization</i>	10
2.3.5 <i>Dropout</i>	12
2.3.6 <i>Stochastic Gradient Descent</i>	13
2.3.7 Pelatihan <i>Deep Neural Network</i>	14
2.3.8 Pengujian <i>Deep Neural Network</i>	17

2.4.	<i>Object Oriented Analysis and Design</i>	17
2.5.	<i>Unified Modeling Language</i>	21
2.5.1	<i>Use Case Diagram</i>	21
2.5.2	<i>Class Diagram</i>	23
2.5.3	<i>Sequence Diagram</i>	24
2.5.4.	<i>Communication Diagram</i>	25
2.5.5.	<i>Deployment Diagram</i>	25
2.6.	Codeigniter.....	26
2.7.	Django.....	26
2.8.	Keras.....	26
BAB III METODOLOGI PENELITIAN.....		27
3.1.	Garis Besar Penyelesaian Masalah.....	27
3.1.1.	Pengumpulan Data.....	28
3.1.2.	Pemetaan Data.....	28
3.1.3.	Pembagian Data Menjadi Data Latih dan Data Uji.....	29
3.1.4.	Arsitektur <i>Deep Neural Network</i>	29
3.1.5.	Pelatihan dan Pengujian <i>Deep Neural Network</i>	30
3.1.5.1.	Pelatihan dan Pengujian <i>Deep Neural Network</i> tanpa <i>Batch Normalization</i>	31
3.1.5.1.1.	Pelatihan <i>Deep Neural Network</i> tanpa <i>Batch</i> <i>Normalization</i>	31
3.1.5.1.2.	Pengujian <i>Deep Neural Network</i> tanpa <i>Batch</i> <i>Normalization</i>	38
3.1.5.2.	Pelatihan dan Pengujian <i>Deep Neural Network</i> dengan <i>Batch Normalization</i>	38
3.1.5.2.1.	Pelatihan <i>Deep Neural Network</i> dengan <i>Batch</i> <i>Normalization</i>	38
3.1.5.2.2.	Pengujian <i>Deep Neural Network</i> dengan <i>Batch</i> <i>Normalization</i>	56
3.1.6.	Evaluasi.....	57
3.2.	Analisis Aplikasi.....	57
3.2.1.	<i>Business Process</i>	58

3.2.2. <i>Requirement</i>	59
3.2.2.1. Deskripsi Umum Aplikasi	60
3.2.2.2. Daftar Aktor	60
3.2.2.3. Daftar <i>Use Case</i>	61
3.2.2.4. <i>Use Case Detail</i>	61
3.2.2.5. <i>Use Case Diagram</i>	63
3.2.2.6. Survei <i>Use Case</i>	65
3.2.2.7. Sketsa Antarmuka	65
3.2.3. <i>Analysis</i>	69
3.2.3.1. Analisis <i>Class Diagram</i>	69
3.2.3.2. <i>Communication Diagram</i>	69
3.2.4. <i>Design</i>	72
3.2.4.1. <i>Sequence Diagram</i>	73
3.2.4.2. <i>Class Diagram</i>	74
3.2.4.3. Skema Basis Data	76
3.2.4.4. <i>Deployment Diagram</i>	78
BAB IV HASIL DAN PEMBAHASAN	80
4.1. Implementasi	80
4.1.1. Spesifikasi Aplikasi	80
4.1.2. Implementasi Basis Data	81
4.1.3. Implementasi <i>Source Code</i>	83
4.1.4. Implementasi Antarmuka	83
4.2. Pengujian Aplikasi	87
4.2.2. Rencana Pengujian Aplikasi	87
4.2.3. Hasil Pengujian Aplikasi	87
4.3. Pengujian Metode <i>Deep Neural Network</i>	88
4.3.2. Skenario Pengujian	88
4.3.2.1. Skenario 1	88
4.3.2.2. Skenario 2	88
4.3.3. Pembahasan Skenario Pengujian	89
4.3.3.1. Pembahasan Skenario 1	89
4.3.3.2. Pembahasan Skenario 2	89

BAB V KESIMPULAN DAN SARAN	92
5.1. Kesimpulan	92
5.2. Saran	92
DAFTAR PUSTAKA.....	93
LAMPIRAN	

DAFTAR GAMBAR

Gambar 2.1. Arsitektur DNN (Karthik, et al., 2018).....	8
Gambar 2.2. Ilustrasi Fungsi Aktivasi ReLU (Agarap, 2018).....	9
Gambar 2.3. Ilustrasi Fungsi Aktivasi <i>Sigmoid</i> (Balaji & Baskaran, 2013).....	9
Gambar 2.4. Model <i>Dropout</i> dalam NN. (a) NN sebelum dilakukan <i>Dropout</i> , (b) NN sesudah dilakukan <i>Dropout</i> (Srivastava, et al., 2014).....	13
Gambar 2.5. <i>Actor</i> (Arlow & Neustadt, 2002)	22
Gambar 2.6. <i>Use Case</i> (Arlow & Neustadt, 2002).....	22
Gambar 2.7. <i>Use Case Diagram</i> (Arlow & Neustadt, 2002)	23
Gambar 2.8. <i>Use Case Detail</i> (Arlow & Neustadt, 2002).....	23
Gambar 2.9. <i>Class Diagram</i> (Docherty, 2005)	24
Gambar 2.10. <i>Sequence Diagram</i> (Docherty, 2005)	24
Gambar 2.11. <i>Communication Diagram</i> (Docherty, 2005).....	25
Gambar 2.12. <i>Deployment Diagram</i> (Docherty, 2005).....	25
Gambar 3.1. Diagram Garis Besar Penyelesaian Masalah	27
Gambar 3.2. Diagram alir proses pemetaan data.....	28
Gambar 3.3. Arsitektur metode DNN tanpa <i>Batch Normalization</i>	29
Gambar 3.4. Contoh penerapan <i>Batch Normalization</i> pada unit 1 pada layer 1 dalam metode DNN.	30
Gambar 3.5. Arsitektur pembahasan metode DNN.....	31
Gambar 3.6. <i>Business Process</i> Aplikasi	59
Gambar 3.7. <i>Use Case Diagram</i> Aplikasi	64
Gambar 3.8. Sketsa Antarmuka Halaman Awal.....	65
Gambar 3.9. Sketsa Antarmuka Halaman Tambah <i>Dataset</i>	66
Gambar 3.10. Sketsa Antarmuka Halaman Daftar <i>Dataset</i>	66
Gambar 3.11. Sketsa Antarmuka Halaman Rincian <i>Dataset</i>	67
Gambar 3.12. Sketsa Antarmuka Halaman Rincian Metode.....	67
Gambar 3.13. Sketsa Antarmuka Halaman Perhitungan.	68
Gambar 3.14. Sketsa Antarmuka Halaman Riwayat Perhitungan.....	68
Gambar 3.15. Sketsa Antarmuka Halaman Rincian Perhitungan.....	69
Gambar 3.16. <i>Class Diagram</i> Fase Analisis.....	70
Gambar 3.17. <i>Communication Diagram</i> Menambah <i>Dataset</i>	72
Gambar 3.18. <i>Communication Diagram</i> Menghapus <i>Dataset</i>	72

Gambar 3.19. <i>Communication Diagram</i> Proses Perhitungan.....	73
Gambar 3.20. <i>Communication Diagram</i> Melihat Riwayat Perhitungan	73
Gambar 3.21. <i>Sequence Diagram</i> Menambah Dataset.....	75
Gambar 3.22. <i>Sequence Diagram</i> Menghapus Dataset	75
Gambar 3.23. <i>Sequence Diagram</i> Proses Perhitungan	75
Gambar 3.24. <i>Sequence Diagram</i> Melihat Riwayat Perhitungan.....	76
Gambar 3.25. <i>Class Diagram</i> Fase Desain.....	76
Gambar 3.26. <i>Deployment Diagram</i> Aplikasi	79
Gambar 4.1. Implementasi Tabel <i>Data_sample</i>	82
Gambar 4.2. Implementasi Tabel <i>Ds_data</i>	82
Gambar 4.3. Implementasi Tabel <i>History</i>	82
Gambar 4.4. Implementasi Antarmuka Halaman Awal.	83
Gambar 4.5. Implementasi Antarmuka Halaman Tambah <i>Dataset</i>	84
Gambar 4.6. Implementasi Antarmuka Halaman Daftar <i>Dataset</i>	84
Gambar 4.7. Implementasi Antarmuka Halaman Rincian <i>Dataset</i>	85
Gambar 4.8. Implementasi Antarmuka Halaman Rincian Metode.	85
Gambar 4.9. Implementasi Antarmuka Halaman Perhitungan.....	86
Gambar 4.10. Implementasi Antarmuka Halaman Riwayat Perhitungan	86
Gambar 4.11. Implementasi Antarmuka Halaman Rincian Perhitungan	87
Gambar 4.12. Grafik Pengaruh <i>Hidden Layer</i> Terhadap Nilai Akurasi DNN-S.....	90
Gambar 4.13. Grafik Pengaruh <i>Hidden Layer</i> Terhadap Nilai Akurasi DNN-BN.	91

DAFTAR TABEL

Tabel 2.1. Daftar Penelitian Menggunakan Metode DNN	6
Tabel 2.2. Fase OOAD (Docherty, 2005).....	18
Tabel 3.1. Tabel Hasil Pemetaan Data	29
Tabel 3.2. Daftar Aktor Aplikasi	60
Tabel 3.3. Daftar <i>Use Case</i> Aplikasi	60
Tabel 3.4. <i>Use Case Detail</i> Menambah <i>Dataset</i>	61
Tabel 3.5. <i>Use Case Detail</i> Menghapus <i>Dataset</i>	62
Tabel 3.6. <i>Use Case Detail</i> Proses Perhitungan	63
Tabel 3.7. <i>Use Case Detail</i> Melihat Riwayat Perhitungan.....	64
Tabel 3.8. Hasil Identifikasi <i>Class Analysis</i>	70
Tabel 3.9. Attribute Tiap <i>Class</i>	71
Tabel 3.10. Rancangan Tabel Data_sample	77
Tabel 3.11. Rancangan Tabel Ds_data	77
Tabel 3.12. Rancangan Tabel Ds_data	78
Tabel 4.1. Hasil Pengujian Black Box Testing.....	88
Tabel 4.2. Hasil pengujian dengan jumlah <i>hidden layer</i> 2-6 pada DNN-S.....	90
Tabel 4.3. Hasil pengujian dengan jumlah <i>hidden layer</i> 2-6 pada DNN-BN.....	91

BAB I

PENDAHULUAN

Bab ini membahas latar belakang masalah, rumusan masalah, tujuan dan manfaat, ruang lingkup, dan sistematika penulisan dalam pembuatan skripsi mengenai Klasifikasi Kanker Payudara dengan Data *MicroRNA* Menggunakan Metode *Deep Neural Network* (DNN) dengan Optimasi *Batch Normalization*.

1.1 Latar Belakang

Kanker payudara atau *Carsinoma Mammae* adalah pertumbuhan sel yang tidak terkendali pada jaringan penunjang payudara yang mengelilingi *lobular* (kelenjar penghasil susu), *ductus* (puting payudara), pembuluh darah dan pembuluh limfe (Society, 2017). Menurut data *International Agency for Research on Cancer* ((IARC), 2018), kanker payudara merupakan penyebab kematian terbanyak di dunia urutan kedua dengan 2.008.849 kasus dan 626.679 di antaranya dinyatakan meninggal dunia. Data statistik kematian di seluruh dunia berdasarkan wilayah *International Agency for Research on Cancer* ((IARC), 2018) menunjukkan bahwa Asia berada pada urutan pertama kasus kanker payudara dengan 911.014 kasus dan 310.577 dinyatakan meninggal dunia. Bahkan, data statistik kanker payudara menurut jenis kelamin *International Agency for Research on Cancer* ((IARC), 2018) juga menunjukkan bahwa jumlah kematian wanita karena kasus kanker payudara menempati urutan pertama di dunia.

Salah satu upaya yang dapat dilakukan untuk mengurangi jumlah kasus kematian akibat kanker adalah melakukan deteksi dini. Menurut Sabrida dalam Buletin Jendela Data & Informasi Kesehatan (2015), deteksi dini kanker tidak hanya dapat menurunkan angka kematian akibat kanker, tetapi juga meningkatkan kualitas hidup penderitanya (RI, 2015). Deteksi dini menggunakan komputer merupakan bantuan otomatis yang memberikan hasil akurat untuk menganalisis suatu penyakit (Karthik, et al., 2018). Klasifikasi merupakan metode yang digunakan oleh mesin untuk memilah atau mengklasifikasikan objek berdasarkan ciri tertentu (Ahmad, 2017).

Salah satu cara yang dapat dilakukan untuk melakukan deteksi dini kanker payudara adalah melakukan klasifikasi pada data *MicroRNA*. *MicroRNAs* (*miRNAs*) adalah kelompok *Ribonucleic Acid* (RNA) nonkode kecil (*ncRNAs*) yang mengatur

ekspresi gen dengan menargetkan RNA pembawa yang sesuai (*mRNA*) (Lan, et al., 2015). *MicroRNA* dapat digunakan untuk mengenali apakah suatu sel merupakan sel kanker atau bukan bahkan dalam stadium paling awal (Madhavan, et al., 2013).

Proses klasifikasi *MicroRNA* dapat dilakukan dengan berbagai cara. Salah satu teknik klasifikasi yang dapat diterapkan adalah dengan menggunakan teknik *Deep Learning*. *Deep Learning* atau juga disebut dengan *Deep Neural Network* (DNN) adalah salah satu teknik pembelajaran mesin terbaru dan paling baik untuk pengenalan pola (Spencer, et al., 2015). DNN memiliki beberapa lapisan tersembunyi (*hidden layer*) dan membuat model ini menjadi model ekspresif yang dapat mempelajari hubungan rumit antara *input* dan *output*-nya (Srivastava, et al., 2014). Lapisan tersembunyi ini secara bertahap akan memetakan fitur tingkat rendah yang tidak terorganisir ke dalam representasi data tingkat tinggi (Spencer, et al., 2015).

Terdapat berbagai penelitian menggunakan DNN. Sebuah penelitian (Thomas & Sael, 2017) untuk memprediksi urutan prekursor *miRNA* menggunakan *dataset miRBase 18.0 release* dengan 58 fitur menunjukkan bahwa DNN dengan akurasi 96.8% mengungguli metode *Support Vector Machine* (SVM) dengan akurasi 93.2%, *Neural Network* (NN) dengan akurasi 91.7%, *Naive Baye's Classifiers* dengan akurasi 91.4%, *K-Nearest Neighbors* dengan akurasi 90.8%, dan *Random Forests* dengan akurasi 93.7%. Penelitian oleh (Waspada, et al., 2017) menerapkan metode *Machine Learning* yang terawasi untuk klasifikasi kanker menurut data *MicroRNA*. Penelitian tersebut menunjukkan bahwa metode *Deep Learning* dengan akurasi 91.49% lebih unggul dari metode *Naive Bayes* dengan akurasi 61.54%, *Decision Tree* dengan akurasi 34.15%, dan *Neural Network* dengan akurasi 5.48%. Penelitian lain oleh (Joshi & Metha, 2018) menggunakan metode *Neural Network* dan *Deep Neural Network* menggunakan data *Wisconsin Breast Cancer Dataset* (WBCD) yang didapatkan dari *UCI repository* untuk klasifikasi penyakit kanker payudara. Penelitian tersebut menunjukkan bahwa *Neural Network* dengan *Linear Discriminant Analysis* (LDA) memberikan hasil yang lebih baik dalam hal akurasi dibandingkan dengan *Deep Neural Network* yaitu (97.06%), spesifitas (100%), dan presisi (100%). *Deep Neural Network* dengan teknik LDA menghasilkan akurasi (62.94%), spesifitas (100%), presisi (100%).

Dari penelitian yang telah dikemukakan sebelumnya, metode DNN memiliki kinerja yang baik, namun metode DNN masih memiliki kelemahan yaitu perubahan

dalam distribusi *input* setiap lapisan yang menyebabkan masalah karena lapisan perlu terus beradaptasi dengan distribusi baru sehingga menghasilkan nilai akurasi yang kurang optimal (Ioffe & Szegedy, 2015). Penelitian (Ioffe & Szegedy, 2015) mengatasi kelemahan ini dengan mengusulkan algoritma *Batch Normalization*. *Batch Normalization* menjadikan normalisasi sebagai bagian dari arsitektur model yang membuat jaringan lebih stabil dengan rentang gradient yang lebih kecil (Ioffe & Szegedy, 2015).

Pada penelitian ini akan dilakukan klasifikasi kanker payudara dengan data *MicroRNA* menggunakan metode DNN tanpa optimasi *Batch Normalization* dan metode DNN dengan optimasi *Batch Normalization*.

1.2 Rumusan Masalah

Berdasarkan latar belakang di atas maka dapat dirumuskan suatu permasalahan yaitu bagaimana perbandingan nilai akurasi klasifikasi kanker payudara berdasarkan data *MicroRNA* menggunakan metode DNN tanpa optimasi *Batch Normalization* dan metode DNN dengan optimasi *Batch Normalization*.

1.3 Tujuan dan Manfaat

Tujuan umum dari penelitian skripsi ini adalah membandingkan nilai akurasi klasifikasi kanker payudara berdasarkan data *MicroRNA* menggunakan metode DNN tanpa optimasi *Batch Normalization* dan metode DNN dengan optimasi *Batch Normalization*.

Manfaat dari penelitian ini adalah agar dapat mengetahui pengaruh *Batch Normalization* terhadap metode *Deep Neural Network* berdasarkan nilai akurasi pada klasifikasi kanker payudara berdasarkan data *MicroRNA*.

1.4 Ruang Lingkup

Dalam penyusunan skripsi ini perlu adanya batasan-batasan agar pembahasan lebih terarah dan tidak melebihi target yang diteliti. Batasan-batasan tersebut adalah:

1. Data yang digunakan sebagai dasar dalam melakukan penelitian ini adalah data ekspresi *MicroRNA* dari sel kanker dan sel normal.
2. Hasil klasifikasi yang akan dikeluarkan yaitu berupa nilai akurasi metode menggunakan masukan data *MicroRNA*.
3. Perhitungan data *MicroRNA* dilakukan menggunakan metode DNN.

4. Optimasi perhitungan DNN menggunakan algoritma *Batch Normalization*.
5. Jumlah *hidden unit*, *dropout rate*, dan *epoch* yang digunakan dalam arsitektur metode DNN adalah 200 unit, 50% dan 100 *epoch*.
6. Data *MicroRNA* akan dibagi dengan perbandingan 70% untuk data latih dan 30% untuk data uji.

1.5 Sistematika Penulisan

Sistematika penulisan yang digunakan dalam skripsi ini terbagi dalam beberapa pokok bahasan, yaitu:

BAB I PENDAHULUAN

Bab ini membahas latar belakang masalah, rumusan masalah, tujuan dan manfaat, ruang lingkup, serta sistematika penulisan dalam penyusunan skripsi.

BAB II LANDASAN TEORI

Bab ini menyajikan hasil studi pustaka mengenai teori yang berhubungan dengan pelaksanaan dan penyusunan skripsi.

BAB III METODOLOGI PENELITIAN

Bab ini menyajikan metodologi penelitian yang digunakan dalam klasifikasi kanker payudara dengan data *MicroRNA* Menggunakan metode *Deep Neural Network* (DNN) tanpa dan dengan optimasi *Batch Normalization*, meliputi garis besar penyelesaian masalah serta analisis dan desain perangkat lunak.

BAB IV HASIL DAN PEMBAHASAN

Bab ini menyajikan deskripsi umum perangkat lunak, analisis kebutuhan perangkat lunak, desain perangkat lunak, implementasi perangkat lunak, dan analisis hasil penelitian mengenai klasifikasi kanker payudara dengan data *MicroRNA* Menggunakan metode *Deep Neural Network* (DNN) tanpa dan dengan optimasi *Batch Normalization*.

BAB V PENUTUP

Bab ini berisi kesimpulan dari uraian yang telah dijabarkan pada bab-bab sebelumnya dan saran untuk pengembangan penelitian lebih lanjut.