

PERBANDINGAN METODE *DEEP LEARNING* DAN *MACHINE LEARNING* UNTUK KLASIFIKASI (UJICOBA PADA DATA PENYAKIT KANKER PAYUDARA)

Yudicy Amelia¹, Puspa Eosina², Foni Agus Setiawan³

Universitas Ibn Khaldun Bogor

Jl. K. H. Sholeh Iskandar Km. 2 Kedung Badak Bogor 16162

Email: yudicyamelia@gmail.com

ABSTRAK

PERBANDINGAN METODE *DEEP LEARNING* DAN *MACHINE LEARNING* UNTUK KLASIFIKASI (UJI COBA PADA DATA PENYAKIT KANKER PAYUDARA). Penyebab kematian utama di seluruh dunia merupakan penyakit kanker, salah satunya adalah kanker payudara. Faktor penentu penyakit tersebut termasuk kategori ganas atau jinak bisa dilihat dari sembilan faktor utama berdasarkan ciri kanker tersebut. Tujuan dari penelitian ini adalah mendapatkan perbandingan metode yang paling akurat terhadap pengklasifikasian penyakit kanker payudara dan melihat selisih nilai *mean square error* (MSE) dari metode *deep learning* dan *machine learning* serta pencocokan hasil klasifikasi kedua metode dengan pernyataan klasifikasi yang sudah ada sehingga didapatkan nilai perbandingan metode yang paling akurat. Data yang digunakan pada penelitian ini diambil dari UCI Edu, dengan data latih sebanyak 546 data dan data uji sebanyak 137 data. Dari proses pengklasifikasian menggunakan FFNN pada *machine learning* dan RBM pada *deep learning*, maka dapat dilihat ada empat (2.92%) data kanker payudara pada proses FFNN yang tidak memenuhi *class* klasifikasi, sedangkan data yang tidak memenuhi *class* klasifikasi pada proses RBM ada dua (1.46%) data. Dalam kasus klasifikasi penyakit kanker payudara, akurasi metode *machine learning* lebih kecil dibandingkan dengan akurasi *deep learning* dan hasil hipotesa menggunakan uji t menunjukkan bahwa nilai signifikansi (*2-tailed*) lebih kecil dari $\alpha = 0.05$ yaitu sebesar $8,68844E-16$ yang menandakan bahwa *deep learning* lebih baik dibandingkan *machine learning*. Dengan akurasi metode *machine learning* sebesar 97.0803% dan *deep learning* sebesar 98.5401%. Nilai MSE pada pengklasifikasian menggunakan FFNN adalah sebesar 0.0814, sedangkan pada RBM sebesar 0.0584

Kata kunci : *deep learning*, *machine learning*, RBM, FFNN, klasifikasi, kanker, kanker payudara

ABSTRACT

COMPARISON OF *DEEP LEARNING* AND *MACHINE LEARNING* METHODS FOR CLASSIFICATION (TESTING ON DISEASE DATA BREAST CANCER). The main cause of death worldwide is cancer, one of which is breast cancer. The determinants of the disease including the category of malignant or benign can be seen from the nine main factors based on the characteristics of the cancer. The purpose of this study was to obtain a comparison of the most accurate method for classifying breast cancer and to see the difference in the mean square error (MSE) of the *deep learning* and *machine learning* methods as well as matching the results of the classification of the two methods with the existing classification statement so that the comparison value of the method was obtained. the most accurate. The data used in this study were taken from UCI Edu, with 546

training data and 137 test data. From the classification process using FFNN in machine learning and RBM in deep learning, it can be seen that there are four (2.92%) breast cancer data in the FFNN process that do not meet the classification class, while the data that does not meet the classification class in the RBM process are two (1.46%) data. In the case of breast cancer classification, the accuracy of the machine learning method is smaller than the accuracy of deep learning and the results of the hypothesis using the t test indicate that the significance value (2-tailed) is smaller than $= 0.05$, which is $8.68844E-16$ which indicates that deep learning is better than machine learning. With machine learning method accuracy of 97,0803% and deep learning of 98,5401%. The MSE value in the classification using FFNN is 0.0814, while the RBM is 0.0584

Keywords: *deep learning, machine learning, RBM, FFNN, classification, cancer, breast cancer*

1. PENDAHULUAN

Representasi buatan dan otak manusia yang selalu mencoba mensimulasikan proses pembelajaran pada otak manusia salah satunya adalah neural network atau jaringan syaraf tiruan. Istilah tiruan atau buatan digunakan karena jaringan syaraf ini diimplementasikan dengan menggunakan program komputer yang mampu menyelesaikan sejumlah proses perhitungan selama proses pembelajaran. Model neural network mengadaptasi metode backpropagation untuk pembelajarannya. Pada awalnya metode backpropagation dirancang untuk neural network feedforward Deep learning melakukan model abstraksi pada data menggunakan algoritma supervised dan unsupervised learning, untuk belajar dari berbagai tingkat abstraksi. Deep learning menggunakan hirarki data untuk klasifikasi, karena peningkatan eksponensial data dalam aplikasi ini deep learning berguna untuk memprediksi keakuratan data. Deep learning banyak digunakan untuk mengambil keputusan, pencarian informasi, dan pengindeksian semantik. Deep learning merupakan

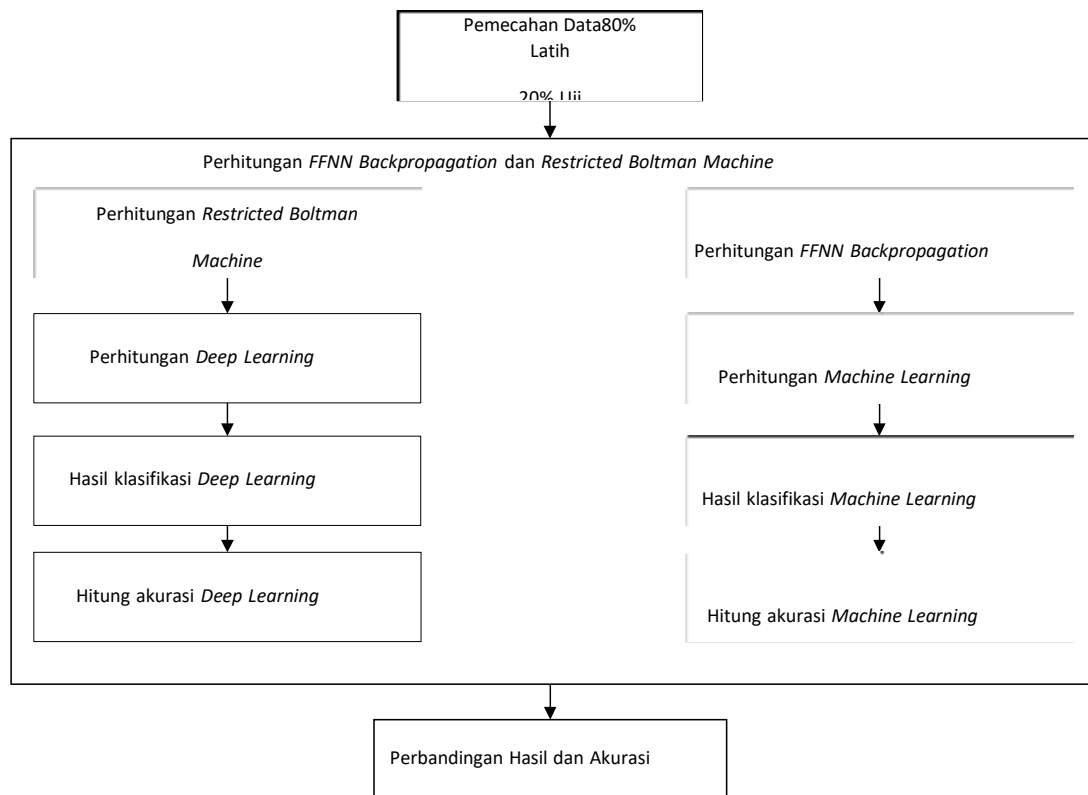
representasi data yang mempunyai beberapa lapisan. Hal ini dapat mengefisienkan proses peningkatan pada data yang jumlahnya besar. Selain itu deep learning cocok digunakan untuk menganalisis data yang tidak terstruktur dan kumpulan data heterogen dari berbagai sumber Metode deep learning mempunyai algoritma yang mampu membantu menemukan parameter pada tiap unit tersembunyi. Deep belief network adalah metode dengan bentuk multi layers neural network Building block pada arsitektur ini disebut dengan Restricted Boltzmann Machines (RBM) yang digunakan untuk merepresentasikan model Penyakit kanker merupakan salah satu penyebab kematian utama di seluruh dunia. Pada tahun 2012, kanker menjadi penyebab kematian sekitar 8,2 juta orang. Kanker paru, hati, perut, kolorektal, dan kanker payudara adalah penyebab terbesar kematian akibat kanker setiap tahunnya. Diketahui bahwa pada tahun 2012 terdapat 14.067.894 kasus baru kanker dan 8.201.575 kematian akibat kanker di seluruh dunia menurut data GLOBOCAN, International Agency for Research on Cancer (IARC) [6]. Resiko kematian

dapat dilihat dari tingkat keganasan kanker, tingkat keganasan kanker pada pasien satu sama lainnya berbeda. Penentuan tingkat keganasan seseorang penderita kanker berdasarkan kondisi kanker

dapat dimodelkan salah satunya menggunakan model klasifikasi. Oleh karena itu data kanker dapat dijadikan sebagai alat uji perbandingan metode deep learning dan machine learning.

2. METODE PENELITIAN

Metode penelitian yang digunakan dalam penyusunan naskah ini meliputi dua bagian pokok yaitu metode pengumpulan data dan metode analisis. Metode penelitian dapat dilihat pada Gambar 1



Gambarang 1 Metode Penelitian

3. HASIL

Hasil uji statistik pada Tabel 1. memperoleh hasil bahwa FFNN pada *machine learning* dan RBM pada *deep learning* dimana perbedaan ini ditunjukkan dari nilai signifikansi (*2-tailed*) yang dihasilkan bernilai lebih kecil dari $\alpha = 0.05$ yaitu sebesar $8,68844E-16$. Hasil ini menunjukkan bahwa H_0 ditolak yang berarti menerima H_1 , yang artinya metode 2 (menggunakan

RBM pada *deep learning*) berbeda nyata dengan metode 1 (menggunakan FFNN pada *machine learning*). Dari hasil akurasi yang dihasilkan terbukti metode 2 memiliki hasil yang lebih baik dari metode 1.

Tabel 1. Hasil Uji T Berpasangan Nilai Akurasi Setiap *Class*

	<i>FFNN</i>	<i>RBM</i>
Mean	2,37425107	2,540145985
Variance	0,55176045	0,794332331
Observations	137	137
Pearson Correlation	0,982430118	
Hypothesized Mean Difference	0	
Df	136	
t Stat	-9,12312283	
P(T<=t) one-tail	4,34422E-16	
t Critical one-tail	1,656134988	
P(T<=t) two-tail	8,68844E-16	
t Critical two-tail	1,977560777	

4. Kesimpulan

Kesimpulan yang didapat setelah melakukan analisis terhadap metode machine learning dan deep learning untuk klasifikasi penyakit kanker payudara adalah sebagai berikut: 1). Hasil klasifikasi dengan FFNN pada machine learning terdapat 4 data yang tidak terklasifikasi dengan benar, sedangkan dengan RBM pada deep learning terdapat 2 data yang tidak terklasifikasi dengan benar. Hasil tersebut didapat dari perbandingan hasil klasifikasi dengan data uji yang berjumlah 137 data kanker payudara. 2). Dalam kasus klasifikasi penyakit kanker payudara, akurasi metode machine learning lebih kecil dibandingkan

dengan akurasi deep learning dan hasil hipotesa menggunakan uji t menunjukkan bahwa nilai signifikansi (2-tailed) lebih kecil dari $\alpha = 0.05$ yaitu sebesar 8,68844E-16 yang menandakan bahwa deep learning lebih baik dibandingkan machine learning. Dengan akurasi metode machine learning sebesar 97.0803% dan deep learning sebesar 98.5401%. Nilai MSE pada pengklasifikasian menggunakan FFNN adalah sebesar 0.0814, sedangkan pada RBM sebesar 0.0584.

5. Daftar Pustaka

[1] Widodo P. Rule-Based Classifier untuk Mendeteksi

- Penyakit Liver. Bianglala Informatika. 2014. Vol II. No 1.
- [2] Rudianto. Penentuan Penyakit Peradangan Hati Dengan Menggunakan Neural Network Backpropagation. Indonesian Journal on Computer Information Technology. 2016. Vol 1. No 1.
- [3] Hirak K dkk. Big Data Analytics In Bio Informatics: Machine Learning Perspective. Jurnal of iatex class files. 2014. Vol 13. No 9.
- [4] Putra BRD, Kusuma WA, dan Kustiyo A. Klasifikasi khasiat Formla Jamu dengan Metode Deep Belief Network. Makalah Kolokium Program S1 Ilmu Komputer Alih Jenis, Departemen Ilmu Komputer, FMIPA-IPB. 2011.
- [5] Ersti Renette, Wisesty Untari Novia, Jonri. Klasifikasi Sinyal EEG Menggunakan Deep Neural Network. Teknik Informatika, Fakultas Teknik, Universitas Telkom.
- [6] Pusat data dan informasi. Situasi Penyakit Kanker. Buletin Jendela Pusat Data dan Informasi. 2015. ISSN 2088 - 270X.
- [7] Susilawati. Algoritma Restricted Boltzmann Machine (RBM) untuk Pengenalan Tulisan Tangan Angka. Seminar Nasional Teknologi Informatika. Universitas Medan Area. 2017. ISBN:978-602-50006-0-7.
- [8] Seonwoo M, Byunghan L, dan Sungroh Y. Deep Larning in Bioinformatics. Department of Electrical and Computer Engineering, Seoul National University, Seoul 08826, Korea.
- [9] Feriante J. Massiveli Multitask Deep Learning for Drug Discovery. Master of science in computer science, university of wisconsin- Madison. 2015.
- [10] Diwandari S dan Setiawan NA. Perbandingan Algoritme J48 dan NBTree untuk Klasifikasi Diaognosa Penyakit pada Soybean. Seminar Nasional Teknologi Informasi dan Komunikasi (SENTIKA). Yogyakarta. 2015. ISSN 2089 – 9815.
- [11] Saputra RA. Komparasi Algoritma Klasifikasi Data Minning untuk Memprediksi Penyakit Tuberculosis (TB): Studi Kasus Puskesmas Karawang Sukabumi. Seminar Nasional Inovasi dan Tren (SNIT), Proceeding SNIT. 2014. Hal 3.
- [12] Fernanda JW. Boosting Neural Network dan Boosting Cart Pada Klasifikasi Diabetes Militus Tipe II. Jurnal Matematika. 2012. Vol 2. No 2.
- [13] Sinta R, Gernowo R, dan Suryono. Rancang Bangun Sistem Peramalan Konsumsi Daya Listrik dengan Artificial Neural Network Backpropagation. Jurnal Sistem Informasi Bisnis. 2013.
- [14] Beale MH, Hagan MT, dan Demuth HB. MATLAB Neural Network Toolbox™ User's Guide. The MathWorks, Inc. https://www.mathworks.com/help/pdf_doc/nnet/nnet_ug.pdf. Diakses 17 Februari 2018.
- [15] Kusumadewi F. Peramalan Harga Emas Menggunakan

Feedforward Neural Network Dengan Algoritma Backpropagation. Skripsi. FMIPA, Matematika, Universitas Negeri Yogyakarta. 2014.

[16] Hrasko R, Pacheco AGC, dan Krohlong RA. Time Series Prediction using Restricted Boltzmann Machine and Backpropagation. Procedia Computer Science. Information Technology and Quantitative Management (ITQM). 2015. Hal.55 ISSN 990 – 999

[17] Ghazali G dan Jondri. Peramalan Saham Menggunakan DBM (Deep Belief Networ). Fakultas Informatika, Universitas Telkom, Bandung.

[18] Hansun Seng. Penerapan WEMA dalam Peramalan Data IHSG. Ultimatics. 2013. Vol V. No 2. ISSN 2085 – 4552.

[19] Solikin R, Jusak, dan Sutomo E. Sistem Pakar Diagnosis Penyakit pada Ayam Petelur Menggunakan Metode Certainty Factor. Jurnal Sistem Informatika. 2014. JSIKA. Vol 3. No 2.

[20] Laxmi GF. Optimasi Pemilihan Threshold dan Operator Fuzzy Local Binary Pattern Menggunakan Multi Objective Genetic Algorithm. Tesis, Institut Pertanian Bogor. 2012.