

PREDIKSI INDEKS PEMBANGUNAN MANUSIA (IPM) BERDASARKAN JUMLAH PENGGUNA INTERNET, PRODUK DOMESTIK REGIONAL BRUTO (PDRB), JUMLAH PEKERJA, JUMLAH PENDUDUK MISKIN DAN POPULASI PENDUDUK PADA TAHUN 2010

PREDICTION OF HUMAN DEVELOPMENT INDEX (HDI) BASED ON NUMBER OF INTERNET USERS, REGIONAL GROSS DOMESTIC PRODUCT (GDP), NUMBER OF EMPLOYEES, NUMBER OF POPULATION IN POVERTY AND NUMBER OF POPULATION IN 2010

Adit Dwi Fajaryanto¹ & Andry Alamsyah²

¹ Prodi S1 Manajemen Bisnis Telekomunikasi dan Informatika, Fakultas Ekonomi Bisnis, Universitas Telkom

² Prodi S1 Manajemen Bisnis Telekomunikasi dan Informatika, Fakultas Ekonomi Bisnis, Universitas Telkom

¹aditdwifajaryanto@gmail.com, ²andry.alamsyah@gmail.com

Abstrak

Indeks Pembangunan Manusia adalah mengukur capaian pembangunan manusia berbasis sejumlah komponen dasar kualitas hidup. Penyediaan data IPM ditujukan sebagai alat perencanaan dan evaluasi kebijakan pemerintah sedangkan indikator IPM menggambarkan keberhasilan target pembangunan pemerintah. Sehingga dapat dikatakan, data IPM menjadi data yang dapat digunakan dalam membuat kebijakan oleh pemerintah. Namun menurut Idris [1], Menteri Perdagangan (Mendag), Thomas Trikasih Lembong menyebut, validitas data masih menjadi persoalan klasik yang menghambat dalam membuat kebijakan, sejumlah instansi pemerintah pun memiliki data yang berbeda sehingga menciptakan kebingungan antar kementerian dalam membuat kebijakan.

Tujuan dari penelitian ini yaitu untuk mengetahui skenario model klasifikasi Indeks Pembangunan Manusia berdasarkan data Jumlah Pengguna Internet, PDRB, Jumlah Pekerja, Jumlah Penduduk Miskin dan Populasi Penduduk pada tahun 2010 menggunakan metode *decision tree* dan *artificial neural network* serta untuk mengetahui tingkat akurasi prediksi Indeks Pembangunan Manusia menggunakan metode *decision tree* dan *artificial neural network*.

Variabel pada penelitian ini adalah jumlah pengguna internet, Produk Domestik Regional Bruto (PDRB), jumlah pekerja, jumlah penduduk miskin dan populasi penduduk pada tahun 2010. Teknik analisis yang digunakan adalah teknik klasifikasi dengan membandingkan metode *Decision Tree Classification* (Pohon Keputusan) dan *Artificial Neural Network* (ANN).

Kata Kunci: Artificial Neural Network; Back Propagation; CART; Data Mining; Decision Tree; IPM.

Abstract

The Human Development Index is a measure of human development achievements based on a number of basic components of quality of life. HDI data provision is intended as a planning tool and evaluation of government policies while the HDI indicators illustrate the success of the government's development targets. That is to say, the data HDI data into data that can be used in the policy making process by the government. However, according to Idris, Minister of Trade (Ministry of Trade), Thomas Trikasih Lembong call, the validity of the data is still the classic issues that hamper policy-making, a number of government agencies also have different data "that creates confusion among ministries in making policy.

The purpose of this study is to determine the model scenarios classification Human Development Index based on data Number of Internet Users, the GDP, Number of Workers, Poverty rates and Population in 2010 using decision tree and artificial neural network methods and to determine the level of accuracy of prediction of the Human Development Index using decision tree and artificial neural network methods.

The variable in this study is the number of Internet users, Gross Domestic Product (GDP), the number of workers, the number of poor people and the population in 2010. The analysis technique used is the technique of classification by comparing the method Decision Tree Classification (Decision Trees) and Artificial Neural network (ANN).

Keywords: Artificial Neural Network; Back Propagation; CART; Data Mining; Decision Tree; HDI.

1. Pendahuluan

Indeks Pembangunan Manusia (IPM) adalah mengukur capaian pembangunan manusia berbasis sejumlah komponen dasar kualitas hidup yang dibangun melalui pendekatan tiga dimensi dasar: dimensi umur panjang dan sehat yang diukur dari data angka harapan hidup, pengetahuan yang diukur dari data gabungan indikator angka melek huruf dan rata-rata lama sekolah, dan kehidupan yang layak yang diukur dari data indikator kemampuan

daya beli masyarakat terhadap sejumlah kebutuhan pokok yang dilihat dari pengeluaran riil perKapita. Namun menurut Pratama dan Al-Shaikh (2012)[2] di era teknologi seperti sekarang dimana kemudahan mengakses internet bisa dijadikan indikator perhitungan Indeks Pembangunan Manusia. Jumlah Pendapatan Regional Bruto juga berperan penting dalam peningkatan daya beli pada masyarakat yang dimana merupakan salah satu indikator penting dalam penghitungan IPM suatu daerah. Pembangunan suatu wilayah memerlukan kontribusi masyarakat dalam membangun lingkungan hidup yang lebih baik, penduduk yang bekerja dianggap mampu memperoleh pendapatan/keuntungan yang dimana dapat mempengaruhi tingkat kemiskinan, bisa dikatakan penurunan jumlah pengangguran berbanding lurus dengan penurunan tingkat kemiskinan. Kemiskinan bisa dipengaruhi banyak faktor, diantaranya disebabkan tingginya pertumbuhan penduduk. Jumlah penduduk yang banyak dapat menyebabkan kemiskinan karena pertumbuhan penduduk berkembang menurut deret ukur sedangkan produksi pangan berkembang menurut deret hitung, sehingga laju pertumbuhan penduduk tidak akan terkejar oleh laju pertumbuhan makanan.

Berdasarkan latar belakang penelitian ini, maka dapat dirumuskan permasalahan penelitian sebagai berikut:

- 1) Bagaimana skenario model klasifikasi Indeks Pembangunan Manusia berdasarkan data Jumlah Pengguna Internet, PDRB, Jumlah Pekerja, Jumlah Penduduk Miskin dan Populasi Penduduk pada tahun 2010 menggunakan Metode Decision tree dan Artificial Neural Network (ANN)?
- 2) Bagaimana tingkat akurasi prediksi Indeks Pembangunan Manusia berdasarkan data Jumlah Pengguna Internet, PDRB, Jumlah Pekerja, Jumlah Penduduk Miskin dan Populasi Penduduk pada tahun 2010 menggunakan Metode Decision Tree dan Artificial Neural Network (ANN)?

2. Dasar Teori dan Kerangka Pemikiran

2.1. Dasar Teori

Big Data menurut Sujana [3] dapat didefinisikan data yang sudah sangat sulit untuk dikoleksi, disimpan dan dikelola maupun dianalisa dengan menggunakan *system database* yang biasa karena volumenya yang terus berlipat. Sathi [4] menambahkan, alasan Big Data berbeda dengan jenis data lain yang dipakai sebelumnya karena terdapat empat buah "V" yang mengkarakterkan data ini, yakni: *volume*, *velocity*, *variety*, dan *veracity*.

Menurut Han *et al* [5] *data mining* adalah proses menemukan pola yang menarik dan pengetahuan dari sejumlah besar data. Sumber data dapat mencakup *database*, gudang data, *web*, repositori informasi lainnya, atau data yang mengalir ke sistem dinamis. Larose dan Larose [6] menambahkan bahwa terdapat enam teknik yang dimiliki data mining berdasarkan tugas yang bisa dilakukan yaitu deskripsi, estimasi, prediksi, klasifikasi, *clustering*, dan asosiasi.

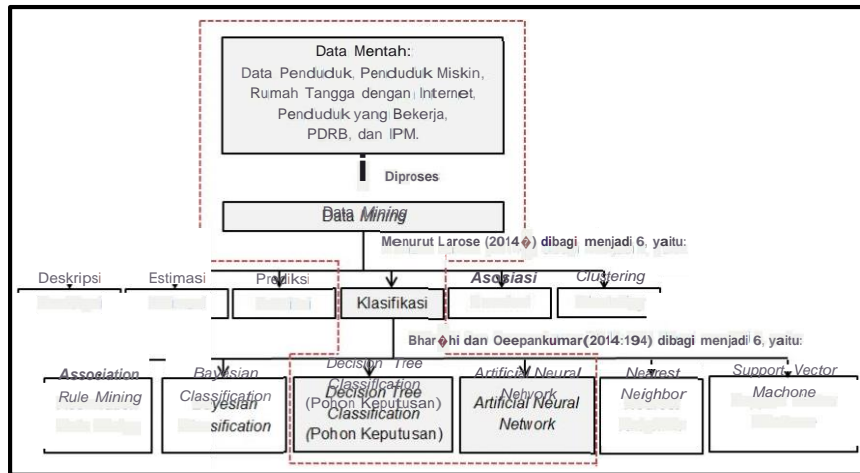
Menurut Mujiasih [7], klasifikasi adalah pengelompokan obyek kedalam satu atau beberapa kelompok berdasarkan variabel yang diamati. Tujuan klasifikasi adalah untuk secara akurat memprediksi kelas target untuk setiap kasus dalam data. Teknik klasifikasi pada data mining dibagi menjadi 2 kelompok menurut Gibert [8] yaitu kelompok untuk teknik klasifikasi yang memiliki variabel respon dan teknik klasifikasi yang tidak memiliki nilai variabel. Menurut Bharathi dan Deepankumar [9] beberapa teknik klasifikasi yang umum digunakan, yaitu: *association rule mining*, *bayesian classification*, *decision tree classification* (pohon keputusan), *nearest neighbor*, *artificial neural networks* (*back propagation*, *support vector machines* (SVMs)).

Decision tree (pohon keputusan) menurut Han *et al* [4] adalah sebuah diagram alir yang mirip dengan struktur pohon, dimana setiap internal *node* menotasikan atribut yang diuji, setiap cabangnya mepresentasikan hasil dari atribut tes tersebut dan *leaf node* mempresentasikan kelas-kelas tertentu atau distribusi dari kelas-kelas. Salah satu algoritma klasifikasi pada *decision tree*, dalam Timofeev [10] adalah *Classification and Regression Tress* (CART).

Artificial Neural Network (ANN) atau juga disebut Jaringan Sistem Tiruan (JST) menurut Suyanto [11], adalah metode pembelajaran yang terinspirasi dari cara kerja sistem syaraf manusia, dimana terdapat neuron atau sel syaraf sebagai unit pemrosesan informasi yang merupakan dasar dari operasi ANN. Adapun nama algoritma dari ANN ini adalah *back propagation*. Sesuai dengan namanya, algoritma ini melakukan dua tahap perhitungan, yaitu: perambatan maju untuk menghitung *error* antara keluaran (*output*) actual dan target; dan perhitungan mundur yang mempropagasikan balik *error* tersebut untuk memperbaiki bobot-bobot pada semua neuron yang ada.

2.2. Kerangka Pemikiran

Kerangka pemikiran penelitian ini digambarkan sebagai berikut:



Gambar 2.1 Kerangka Pemikiran

2.3. Sumber Data

Data yang digunakan pada penelitian ini adalah data sekunder. Menurut Iskandar [12] data sekunder dapat diperoleh dari penelaahan dokumen-dokumen yang berhubungan dengan persoalan penelitian, seperti literatur yang memiliki relevansi dengan fokus kajian. Selain itu data sekunder juga dapat diperoleh melalui arsip perusahaan, website, internet, dan sebagainya. Data yang digunakan disediakan dalam website Badan Pusat Statistik (www.bps.go.id).

3. Pembahasan

3.1. Hasil Analisis dengan Metode Decision Tree

Model klasifikasi Indeks Pembangunan Manusia (IPM) berdasarkan Populasi Penduduk, Penduduk Miskin, Jumlah Pekerja, Pengguna Internet, PDRB pada Tahun 2010 dengan metode decision tree yaitu dapat dilihat pada Gambar 3.1 sebagai berikut:



Gambar 3.1 Model Decision Tree

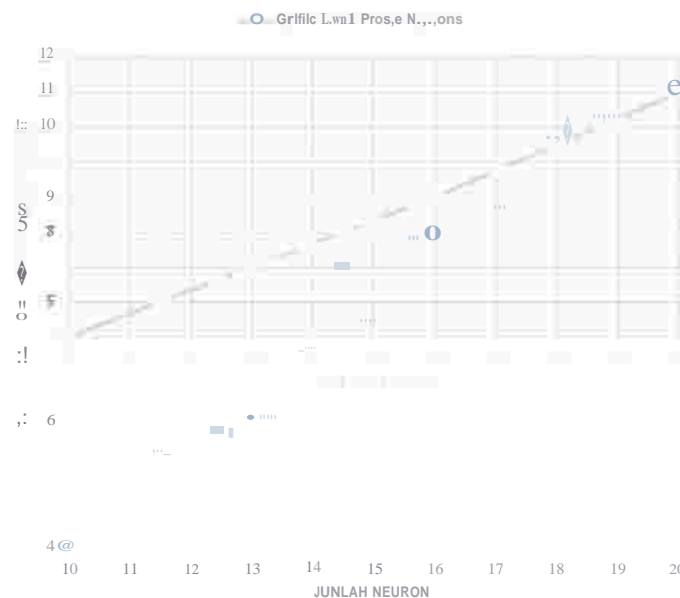
3.2. Hasil Analisis dengan Metode *Artificial Neural Network*

Model terbaik Metode Artificial Neural Network didapatkan dengan melakukan variasi jumlah neuron pada hidden layer dengan variasi 10, variasi 13, variasi 16, dan variasi 20. Setiap variasi selanjutnya diiterasi sebanyak 10 kali untuk mendapatkan model dengan rata-rata error terkecil, model dengan rata-rata error terkecil merupakan model klasifikasi terbaik. Adapun hasil iterasi dengan variasi jumlah neuron dapat dilihat pada Tabel 3.1 berikut :

Tabel 3.1 Hasil Iterasi dengan Variasi Jumlah Neuron

Iterasi Ke-	Jumlah Variasi Neuron							
	10		13		16		20	
	Error	Steps	Error	Steps	Error	Steps	Error	Steps
1	9.642052	61564	7.312040	294549	6.213493	181226	5.231668	217302
2	6.311599	203406	14.807025	90588	8.72494	254536	7.020638	398007
3	8.578704	117722	11.107471	120830	6.213493	187226	12.91957	22395
4	13.1284	56536	9.837222	186449	6.62392	188503	7.813179	279577
5	10.78088	98851	15.280870	64973	9.374675	96564	6.864372	421925
6	9.465707	85662	15.167588	47126	8.735982	107519	7.686853	154735
7	9.890299	173896	12.881257	42094	6.242786	169779	9.114775	196836
8	10.75404	63405	11.874337	166376	11.68992	190169	7.388973	200142
9	11.64513	71811	11.781089	93808	9.156647	75282	7.032150	171096
10	15.11769	67231	11.158436	277420	10.83337	107507	5.523824	138147
Rata-rata	10.53145	100008.4	12.12073	138421.3	8.380923	155831.1	7.6596	220016.2

Tabel 3.1 menjelaskan bahwa nilai rata-rata error dari jumlah neuron 10 adalah 10.53145, untuk jumlah neuron 13 adalah 12.12073, untuk jumlah neuron 16 adalah 8.380923 dan untuk jumlah neuron 20 adalah 7.6596. Selanjutnya, rata-rata jumlah steps dari jumlah neuron 10 adalah 100008.4 dibulatkan menjadi 100008, untuk jumlah neuron 13 adalah 138421.3 dibulatkan menjadi 138421 untuk jumlah neuron 16 adalah 155831.1 dibulatkan menjadi 155831 dan untuk jumlah neuron 20 adalah 220016.2 dibulatkan menjadi 220016. Dapat diketahui dari table 4.2 bahwa nilai rata-rata error terkecil sebesar 7.6596 adalah metode dengan jumlah neuron terbesar yaitu 20. Hal ini berbanding terbalik dengan rata-rata jumlah steps-nya, model dengan jumlah neuron 20 mempunyai jumlah steps terbanyak, yaitu sebesar 220016. Adapun waktu yang dibutuhkan untuk melakukan 1 kali percobaan untuk masing-masing jumlah variasi neuron dapat dilihat pada Gambar 3.2 sebagai berikut:

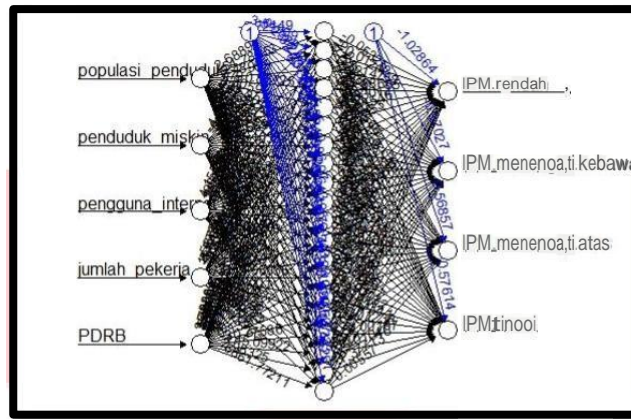


Gambar 3.2 Grafik Lama Proses Neurons

Berdasarkan Gambar 3.2 dapat diketahui untuk tiap satu kali percobaan terhadap variasi jumlah neuron dibutuhkan waktu yang berbeda-beda oleh R-Studio untuk menghasilkan model klasifikasinya, semisal untuk variasi jumlah neuron 16 dibutuhkan waktu sekitar 8 menit untuk satu kali percobaan dan untuk jumlah neuron 20

dibutuhkan waktu sekitar 11 menit untuk satu kali percobaannya. Artinya, semakin banyak jumlah neuron, maka total waktu yang dibutuhkan semakin lama, lalu jumlah steps-nya akan semakin besar, tetapi nilai error-nya akan semakin kecil.

Maka dapat disimpulkan model ANN terbaik adalah 5 : 20 : 4, artinya model terdiri dari lima (5) input layer, 1 hidden layer dengan 20 neuron dan 4 output layer. Adapun model klasifikasi Indeks Pembangunan Manusia (IPM) berdasarkan Jumlah Pengguna Internet, Produk Domestik Regional Bruto (PDRB), Jumlah Pekerja, Jumlah Penduduk Miskin dan Populasi Penduduk pada Tahun 2010 dengan metode artificial neural network yaitu dapat dilihat pada Gambar 3.3 berikut:



Gambar 3.3 Model Klasifikasi dengan Artificial Neural Network

3.3. Hasil Prediksi

Berikut hasil prediksi dari 2 metode klasifikasi Indeks Pembangunan Manusia (IPM), yaitu sebagai berikut:

3.3.1. Hasil Prediksi dengan Metode Decision Tree

Tabel 3.2 Hasil Prediksi Model Decision Tree

	IPM Menengah Atas	IPM Menengah Bawah	IPM Rendah
IPM Menengah Atas	84	0	1
IPM Menengah Bawah	11	1	0
IPM Rendah	1	0	1

Tabel 3.2 merupakan hasil prediksi dari model decision tree menggunakan data test sebesar 20% dari keseluruhan data yang berjumlah 99 data. Gambar 4.4 menjelaskan terdapat 84 kabupaten/kota hasil prediksi tepat yang termasuk dalam kategori IPM Menengah Atas, akan tetapi terdapat 1 kabupaten/kota yang termasuk dalam kategori IPM Rendah namun diprediksikan termasuk dalam kategori IPM Menengah Atas.

Kemudian terdapat 11 kabupaten/kota yang diprediksi termasuk IPM Menengah Bawah dengan namun sebenarnya termasuk dalam katrgoti IPM Menengah Atas, lalu ada 1 kota/kabupaten yang diprediksikan tepat termasuk dalam kategori IPM Menengah Bawah.

Selanjutnya terdapat 1 kabupaten/kota yang termasuk IPM Menengah Atas tetapi diprediksi masuk IPM Rendah dan terdapat 1 kabupaten/kota hasil prediksi tepat termasuk IPM Rendah. Dari 99 data yang diprediksi terdapat 13 misklasifikasi data, sehingga dapat disimpulkan dengan model decision tree kesalahan di prediksi sebesar 13,13 %.

3.3.2. Hasil Prediksi dengan Metode Artificial Neural Network (ANN)

Setelah didapatkan model ANN terbaik, selanjutnya model ANN digunakan untuk memprediksi kabupaten/kota yang termasuk ke dalam kategori IPM berdasarkan lima (5) atribut *input*/prediktornya. Data yang digunakan adalah *data test* sebesar 20% dari data keseluruhan yang dipilih secara *random*. Data ini merupakan

data yang tidak digunakan dalam pembentukan model ANN terbaik. Berikut ini adalah hasil prediksi klasifikasi yang dilakukan dengan menggunakan metode ANN beserta kesalahan klasifikasinya yaitu dapat dilihat pada Tabel 3.3 berikut:

Tabel 3.3 Hasil Prediksi Model *Artificial Neural Network*

	IPM Menengah Atas	IPM Menengah Bawah	IPM Rendah
IPM Menengah Atas	88	1	0
IPM Menengah Bawah	7	2	0
IPM Rendah	1	0	1

Tabel 3.3 merupakan hasil prediksi dari model ANN menggunakan data test sebesar 20% dari keseluruhan data yang berjumlah 99 data. Tabel 4.3 menjelaskan terdapat 88 kabupaten/kota hasil prediksi tepat yang termasuk dalam kategori IPM Menengah Atas, akan tetapi terdapat 1 kabupaten/kota yang sebenarnya termasuk dalam kategori IPM Menengah Bawah namun diprediksi termasuk IPM Menengah Atas.

Kemudian terdapat 7 kabupaten/kota yang termasuk IPM Menengah Atas tetapi diprediksi termasuk IPM Menengah Bawah lalu terdapat 2 kabupaten/kota yang diprediksi termasuk IPM Menengah Bawah dengan tepat.

Selanjutnya terdapat 1 kabupaten/kota yang termasuk IPM Menengah Bawah akan tetapi diprediksian termasuk IPM Rendah, dan terdapat 1 kabupaten/kota yang diprediksian tepat termasuk IPM Rendah. Dari 99 data yang diprediksi terdapat 9 misklasifikasi data, sehingga dapat disimpulkan model ANN 5:20:4 mempunyai kesalahan prediksi sebesar 9,09%

Dari hasil prediksi dengan data test sebesar 20% dari keseluruhan data yang berjumlah 99 data menggunakan 2 metode klasifikasi berbeda yaitu decision tree dan artificial neural network didapatkan tingkat akurasi untuk masing-masing metode. Untuk metode decision tree terdapat 13 misklasifikasi data sehingga tingkat akurasi untuk metode ini sebesar 13,13% sedangkan untuk metode artificial neural network terdapat 9 misklasifikasi data sehingga tingkat akurasi untuk metode ini sebesar 9,09%. Dapat disimpulkan bahwa metode klasifikasi ANN lebih baik dibandingkan model klasifikasi decision tree.

4. Kesimpulan dan Rekomendasi

4.1. Kesimpulan

Berdasarkan hasil penelitian yang telah dilakukan mengenai prediksi Indeks Pembangunan Manusia (IPM) berdasarkan jumlah pengguna internet, Produk Domestik Regional Bruto (PDRB), jumlah pekerja, jumlah penduduk miskin dan populasi penduduk pada tahun 2010, maka dapat ditarik kesimpulan sebagai berikut :

- 1) Model klasifikasi decision tree dibentuk dengan 80% dari keseluruhan data, yaitu berjumlah 398 data. Adapun hasil model prediksi dengan algoritma Classification and Regression Tree (CART) berbentuk pohon keputusan yaitu memiliki 12 node. Kemudian untuk model klasifikasi Artificial Neural Network juga dibentuk dengan 80% dari keseluruhan data, yaitu berjumlah 398 data, model terbaik didapatkan dengan melakukan percobaan (iterasi) sebanyak 10 kali dengan jumlah neuron yang berbeda, yaitu 10, 13, 16 dan 20. Rata-rata error yang paling rendah dipilih sebagai model klasifikasi ANN. Adapun model klasifikasi dengan algoritma back propagation yang dipilih yaitu terdiri dari 5 input layers, 1 hidden layers dengan 20 neuron dan 4 output layers.
- 2) Dari 99 data atau 20% dari keseluruhan data yang di prediksi, terdapat 13 misklasifikasi data sehingga tingkat kesalahan model klasifikasi decision tree sebesar 13,13%, sedangkan untuk model klasifikasi ANN, dari 99 data yang di prediksi, hanya terdapat 9 misklasifikasi data, sehingga kesalahan model ANN hanya sebesar

9,09%. Dapat disimpulkan bahwa metode klasifikasi dengan model ANN lebih baik dibandingkan metode klasifikasi decision tree.

4.2. Rekomendasi

Hasil penelitian ini menemukan bahwa berdasarkan kesimpulan yang telah diuraikan, penulis mengajukan beberapa saran. Saran ini diharapkan dapat menjadi bahan pertimbangan dan perbaikan baik bagi objek penelitian dan peneliti selanjutnya.

4.2.1. Saran untuk Pemerintah

Berdasarkan hasil penelitian yang telah dilakukan, diajukan saran-saran konstruktif untuk Pemerintah Indonesia sebagai objek penelitian. Berikut adalah saran-saran yang diajukan yaitu melakukan *update* data secara berkala, sehingga dapat memudahkan pemerintah dalam membuat dan mengambil kebijakan terkait dengan Indeks Pembangunan Manusia (IPM).

4.2.2. Saran untuk penelitian selanjutnya

Karena penelitian ini hanya dibatasi pada Jumlah Pengguna Internet, Produk Domestik Regional Bruto (PDRB), Jumlah Pekerja, Jumlah Penduduk Miskin dan Populasi Penduduk Pada tahun 2010, maka bagi peneliti selanjutnya disarankan:

- 1) untuk menggunakan parameter-parameter lain dalam menyusun model klasifikasi seperti besaran partisi dataset, penggunaan algoritma lain, dan memperbanyak jumlah neuron.
- 2) untuk menggunakan metode klasifikasi yang lain seperti Association Rule Mining, Bayesian Classification, Nearest Neighbor, dan Support Vector Machines (SVMs).

Daftar Pustaka:

- [1] Idris, Muhammad. (2016). *Mendag Sebut Banyak Kebijakan Gagal Akibat Data Berantakan*. [Online]. <http://finance.detik.com/read/2016/01/27/134707/3128357/4/mendag-sebut-banyak-kebijakan-gagal-akibat-data-berantakan> [Diakses 24 Maret 2016]
- [2] Pratama, Ahmad Raf'ie dan Al-Shaik, Moneer. (2012). Relation and Growth of Internet Penetration Rate with Human Development Level from 2000 to 2010 (Volume 2012). *Jurnal pada Communications of the IBIMA*.
- [3] Sujana, Aprianti Putri. (2013). *Memanfaatkan Big Data untuk Mendeteksi Emosi* – *Jurnal Teknik Komputer Unikom*, 2 (2), 1-4. Retrieved from Academia.
- [4] Sathi, Arvind. (2012). *Big Data Analytics: Disruptive Technologies for Changing the Game*. Boise: MC Press.
- [5] Han *et al.* (2012). *Data Mining Concepts and Techniques Third Edition*. Waltham: Elsevier.
- [6] Larose, Daniel T., dan Larose, Chantal D. (2014). *Discovering Knowledge in Data: An Introduction to Data Mining Second Edition*. New Jersey: John Wiley & Sons Inc.
- [7] Mujiasih, Subekti. (2011). *Pemanfaatan Data Mining untuk Prakiraan Cuaca* - *Jurnal Meteorologi dan Geofisika*, 12(2), 189-195. Retrieved from Badan Meteorologi, Klimatologi, dan Geofisika Pusat Penelitian dan Pengembangan.
- [8] Gibert *et al.* (2010). Choosing the Right Data Mining Technique: Classification of Methods and Intelligent Recommendation – International Congress on Environmental Modelling and Software Modelling for Environment's Sake, 1-9. Retrieved from <http://www.iemss.org/>.
- [9] Bharathi, A., & Deepankumar, E. (2014). Survey on Classification Techniques in Data Mining - International Journal of Recent and Innovation Trends in Computing and Communication (IJRITCC), 2(7), 1983-1986. Retrieved from <http://www.ijritcc.org>.
- [10] Timofeev, Roman. (2004). *Classification and Regression Trees (CART) Theory and Applications*. Tesis pada Humboldt University: tidak diterbitkan.
- [11] Suyanto. (2007). *Artificial Intelligence: Searching, Reasoning, Planning, dan Learning*. Bandung: Informatika Bandung.
- [12] Iskandar. (2008). *Metodologi Penelitian Pendidikan dan Sosial (Kuantitatif dan Kualitatif)*. Jakarta: Gaung Persada Press