

## Prediksi Indeks Pembangunan Manusia Menggunakan Algoritma C4.5 di Kabupaten Kampar

Ridwan Miftahul Putra<sup>1</sup>, Elvira Asril<sup>2</sup>, Taslim<sup>3</sup>

<sup>1,3</sup> Program Studi Teknik Informatika Fakultas Ilmu Komputer Universitas Lancang Kuning

<sup>2</sup> Program Studi Sistem Informasi Fakultas Ilmu Komputer Universitas Lancang Kuning

Jl. Yos Sudarso KM. 8 Rumbai, Pekanbaru, Riau, telp. 0811 753 2015

e-mail: <sup>1</sup>ridwanmiftahul12@gmail.com, <sup>2</sup>taslim@unilak.ac.id, <sup>3</sup>elvira@unilak.ac.id

### Abstract

*Indek Pembangunan Manusia merupakan alat ukur dari perbandingan harapan hidup, kemampuan baca tulis, tingkat pendidikan dan standar hidup. Indek Pembangunan Manusia merupakan sebuah tolak ukur untuk mengklasifikasikan sebuah negara. dan juga digunakan sebagai bahan kebijakan ekonomi terhadap kualitas hidup. Dalam penelitian ini menggunakan Algoritma C4.5 untuk memprediksi Indek Pembangunan Manusia di Kabupaten Kampar Riau dan hasilnya dapat digunakan sebagai salah satu alat bantu bagi pihak yang berkepentingan dalam pengambilan kebijakan.. Variabel yang digunakan yaitu harapan hidup, rata-rata lama sekolah, harapan lama sekolah, pengeluaran perkapita dan pengaturan range Level dikelompokkan atas tiga kelompok yaitu tinggi, sedang dan bawah. Pada penelitian ini digunakan data Indek Pembangunan Manusia tahun 2010 sampai dengan tahun 2014. Berdasarkan pengujian yang dilakukan setiap tahun nilai kategori menunjukkan adanya peningkatan, nilai entropy IPM 0,72193 dan gain 0,72193.*

**Keyword :** Indeks Pembangunan Manusia, prediksi, kebijakan,, C4.5

### Abstract

*Human Development Index is a measure of the comparison of life expectancy, literacy, education level and standard of living. The Human Development Index is a benchmark for classifying a country. and also used as an economic policy material for quality of life. In this study using C4.5 Algoritma to predict the Human Development Index in Kampar Regency Riau and the results can be used as one of the tools for stakeholders in policy making ... Variables used are life expectancy, average length of school, long expectations school, per capita expenditure and range level settings are grouped into three groups, namely high, medium and lower. In this study, Human Development Index data from 2010 to 2014 were used. Based on the tests conducted each year the value of the category showed an increase, the IPM entropy value was 0.72193 and the gain was 0.72193.*

**Keyword:** Human Development Index, prediction, policy, C4.5

### 1. Pendahuluan

Indek Pembangunan Manusia merupakan alat ukur yang mempengaruhi kebijakan ekonomi terhadap kualitas hidup [1]. Alat ukur yang digunakan yaitu perbandingan harapan hidup, kemampuan baca tulis, tingkat pendidikan dan standar hidup. Adapun pencapaian pembangunan manusia di bangun melalui pendekatan tiga dimensi dasar yaitu dimensi sehat yang diukur dari data angka harapan hidup, dimensi pengetahuan yang diukur dari angka melek

---

huruf dan rata-rata lama sekolah, dan kehidupan layak yang diukur dari data kemampuan pembelian terhadap sejumlah kebutuhan pokok [2]

IPM ditujukan sebagai alat perencanaan dan evaluasi kebijakan pemerintah, seperti alokasi dana untuk daerah, sedangkan indikator IPM menggambarkan keberhasilan target pembangunan pemerintah, sehingga dapat dikatakan, IPM menjadi data yang dapat digunakan dalam membuat kebijakan oleh pemerintah. Pada penelitian ini digunakan data dari Pusat Badan Statistik Kabupaten Kampar dari tahun 2010 sampai dengan tahun 2014. Hasil dari penelitian ini diharapkan dapat menjadi salah satu alat bantu bagi pihak terkait dalam pengambilan kebijakan terkait peningkatan Indeks Pembangunan Manusia.

Beberapa penelitian terkait IPM sebelumnya juga sudah dilakukan oleh para peneliti. Adit pada tahun 2016 menggunakan algoritma artificial neural network melakukan penelitian prediksi indeks pembangunan manusia berdasarkan jumlah pengguna internet, produk domestik regional bruto, jumlah pekerja, jumlah penduduk miskin dan populasi penduduk pada tahun 2010[3]. Dwi Maumere Putra dan Vita Ratnasari melakukan prediksi IPM provinsi Jawa Timur menggunakan metode regresi logistik ridge dengan menggunakan tiga variabel, yaitu angka kematian bayi, angka buta huruf, dan angka partisipasi sekolah dengan metode backward elimination[4]. Sementara itu Betha juga melakukan penelitian untuk prediksi IPM Jawa Barat dengan menggunakan algoritma dynamic bayesian networks dan mengukur keterkaitan antar variabel yang terlibat didalamnya[5].

Pada penelitian ini digunakan algoritma C4.5 untuk melakukan prediksi IPM pada kabupaten Kampar dengan menganalisis sejumlah atribut yang terkait dengan nilai IPM yaitu harapan hidup, kemampuan tulis baca, tingkat pendidikan dan standar hidup. Algoritma C4.5 ini bekerja dengan cara mempertimbangkan hal-hal yang dapat memberikan dampak pada pilihan keputusan, beserta perkiraan hasil akhir yang akan didapat bila keputusan tersebut diambil [6]. Variabel predictor yaitu angka harapan hidup, harapan lama sekolah, indeks pembangunan manusia, pengeluaran perkapita dan rata-rata lama sekolah.

## 2. Metode Penelitian

Dalam penelitian ini menggunakan metode deskriptif yang bertujuan untuk memecahkan masalah yang ada pada saat ini dan bersifat aktual. Data yang akan diteliti dikumpul, disusun dan dianalisa terlebih dahulu.

### a. Tahap Perencanaan

Adalah tahap yang direncanakan akan dikerjakan dalam proses penelitian ini yaitu

#### 1. Mendefinisikan Masalah

Mengamati dan mencari permasalahan yang akan dibahas pada penelitian, yaitu bagaimana agar metode ini dapat diterapkan dalam tempat penelitian dan penulis dapat memprediksi IPM.

#### Penentuan Tujuan

Tujuan dari penelitian ini adalah untuk memprediksi Indeks Pembangunan Manusia di Badan Pusat Statistik Kampar.

#### 2. Studi Pustaka

Bertujuan untuk mengetahui teori-teori apa yang akan digunakan untuk menyelesaikan permasalahan yang akan diteliti, serta mendapatkan dasar-dasar referensi yang kuat dalam penelitian

#### 3. Tahap Implementasi

Pada tahap ini, hasil dari analisis dan perancangan akan diimplementasikan melalui proses pengkodean sistem.

---

### 3. Hasil Dan Pembahasan

#### a. Decision Tree

*Decision Tree* memgolah data dengan hasil akhir berupa sebuah keputusan yang dapat berfungsi untuk mengeksplorasi hubungan yang antara sejumlah variabel masukan dengan variabel tujuan. Variabel masukan yang digunakan yaitu harapan hidup, rata-rata lama sekolah, harapan lama sekolah, pengeluaran perkapita dengan variable tujuan *low, medium dan high*.

#### b. Algoritma C4.5

Algoritma decision tree C4.5 bekerja dengan tahapan tahapan sebagai berikut [7]

1. Pilih nilai attribute dasar
2. Untuk setiap attribute A, cari gaint rasio informasi yang dinormalisasi dari pemisahan pada A.
3. Buat simpul keputusan
4. Ulangi kembali pada daftar yang diperoleh dengan memisahkan pada a\_best, dan tambahkan simpul tersebut sebagai anak-anak simpul.

Tahap awal dari penelitian ini adalah menentukan data yang akan digunakan untuk menyelesaikan masalah penelitian. Data yang digunakan adalah data IPM Kabupaten Kampar dari tahun 2010 sampai tahun 2014. Adapun data yang digunakan dapat dilihat pada tabel 2.1 dibawah ini :

Tabel 2.1 IPM

Tahun	Harapan Hidup	Rata-rata Lama Sekolah	Harapan Lama Sekolah	Pengeluaran Perkapittak Yang disesuaikan	IPM (BPS)
2010	69,58	8,08	11,65	10,433	68.62
2011	69,65	8,56	11,93	10,513	69.64
2012	69,72	8,59	12,22	10,593	70.08
2013	69,77	8,62	12,51	10,624	70.46
2014	69,80	8,62	12,72	10,651	70.72

Dari data data yang ada selanjutnya diproses untuk menghitung nilai maksimum dan minimum untuk masing masing komponen IPM. Perhitungan mendapatkan nilai IPM pada tahun 2010 dapat dilihat pada table 2.2 dibawah ini :

Tabel 2.2 Nilai Maksimum dan Minimum Komponen IPM

Indikator	Komponen	Nilai Maksimum	Nilai Minimum
Angka Harapan Hidup		85	25
Harapan Lama Sekolah		100	0
Rata-rata Lama Sekolah		15	0
Pengeluaran Perkapital		732,72	300

Indeks angka harapan hidup	: $(69,58 - 25) / (85 - 25) = 0,78210$
Harapan lama sekolah	: $(11,65 - 0) / (100 - 0) = 0,1165$
Indeks rata-rata lama sekolah	: $(8,08 - 0) / (15 - 0) = 0,53866$
Indeks pendidikan	: $2/3 (0,1165) + 1/3 (0,53866) = 0,27636$
Indeks Konsumsi per kapita	: $(10,433 - 300) / (732,72 - 300) = -0,66867$

Angka IPM dapat dihitung menggunakan persamaan :

$$\begin{aligned} \text{IPM} &: 1/3 (\text{Indeks angka harapan hidup} + \text{Indeks pendidikan} + \text{Indeks konsumsi per kapita}) \\ &: 1/3 (0,78210 + 0,27636 + -0,66867) = 0,11812 \\ \text{IPM dalam ratusan} &: 0,11812 * 100 = 11,8116 \end{aligned}$$

Perhitungan mendapatkan IPM Level

$$\begin{aligned} \text{Hasil} &= (\text{Nilai\_tertinggi} - \text{Nilai\_terendah})/3 \\ &= (0,11812 - 0,11035)/3 \\ &= 0,00259 \end{aligned}$$

$$\begin{aligned} \text{High} &= \text{Nilai tertinggi} - \text{Hasil} \\ &= 0,11812 - 0,00259 \\ &= 0,11553 \end{aligned}$$

$$\begin{aligned} \text{Medium} &= \text{High} - \text{Hasil} \\ &= 0,11553 - 0,00259 \\ &= 0,11294 \end{aligned}$$

$$\begin{aligned} \text{Low} &= \text{Medium} - \text{Hasil} \\ &= 0,11294 - 0,00259 \\ &= 0,11035 \end{aligned}$$

Perhitungan mendapatkan kategori

$$\begin{aligned} \text{Hasil} &= (\text{Nilai\_tertinggi} - \text{Nilai\_terendah})/2 \\ &= (0,11812 - 0,11035)/3 \\ &= 0,003885 \end{aligned}$$

$$\begin{aligned} \text{Good} &= \text{Nilai tertinggi} - \text{Hasil} \\ &= 0,11812 - 0,003885 \\ &= 0,114235 \end{aligned}$$

$$\begin{aligned} \text{Bad} &= \text{Good} - \text{Hasil} \\ &= 0,114235 - 0,003885 \\ &= 0,11035 \end{aligned}$$

Berdasarkan IPM Level dan perhitungan kategori maka didapat tabel analisa IPM level dan Kategori data yang akan digunakan untuk tahap perhitungan selanjutnya. Tabel analisa dapat dilihat pada tabel 2.3 berikut

Tabel 2.3 Hasil Analisa IPM Level dan Kategori

Tahun	IPM	IPM Level	Kategori
2010	0.11035	Low	Bad
2011	0.11498	Medium	Good
2012	0.11629	High	Good
2013	0.11746	High	Good
2014	0.11812	High	Good

Dari hasil Analisa IPM level dan kategori dihitung berapa banyak data yang termasuk dalam kategori yang sudah ditentukan, adapun hasil analisa dapat dilihat pada tabel 2.6 berikut.

Tabel 2.4 Jumlah IPM Level dan Kategori

Good	4
Bad	1
Total	5

High	3
Medium	1
Low	1
High Good	3
Total High	3
High Bad	0
Total Bad	1
Medium Good	1
Total Medium	1
Medium Bad	0
Total Bad	1
Low Good	0
Total Low	1
Low Bad	1
Total bad	1

Penyelesaian :

$$Entropy (s) = \sum_{i=1}^n -p_i \log_2 p_i$$

Keterangan:

S=Himpunan kasus

A=Fitur

n=Jumlah partisi S

pi=proporsi  $S_i$  terhadap S

$$= \frac{-Good}{Total} * \log_2 \left( \frac{Good}{Total} \right) + \frac{-Bad}{Total} * \log_2 \left( \frac{Bad}{Total} \right)$$

$$= \frac{-4}{5} * \log_2 \left( \frac{4}{5} \right) + \frac{-1}{5} * \log_2 \left( \frac{1}{5} \right)$$

$$= -0,8 * \log_2(0,8) + -0,2 * \log_2(0,2)$$

$$= -0,8 * -0,321928095 + -0,2 * -2,321928095$$

$$= 0,257542476 + 0,464385619$$

$$= 0,72193$$

Nilai Gain High

$$= \frac{High}{Total} * \left( \frac{-High Good}{Total High} * \log_2 \frac{High Good}{Total High} \right) + \frac{-High Bad}{Total Bad} * \log_2 \left( \frac{High Bad}{Total Bad} \right)$$

$$= \frac{3}{5} * \left( \frac{-3}{3} * \log_2 * \frac{-3}{3} \right) + \frac{0}{1} * \log_2 \left( \frac{0}{1} \right)$$

$$= \frac{3}{5} * (-1 * 0)$$

$$= 0,6 * 0$$

$$= 0$$

Nilai Gain Medium

$$= \frac{Medium}{Total} * \left( \frac{-Medium Good}{Total Medium} * \log_2 \frac{Medium Good}{Total Medium} \right) + \frac{-Medium Bad}{Total Bad} * \log_2 \left( \frac{Medium Bad}{Total Bad} \right)$$

$$= \frac{1}{5} * \left( \frac{-1}{1} * \log_2 * \frac{1}{1} \right) + \frac{0}{1} * \log_2 \left( \frac{0}{1} \right)$$

$$= \frac{1}{5} * (-1 * 0)$$

$$= 0,2 * 0$$

$$= 0$$

Nilai Gain Low

$$= \frac{Low}{Total} * \left( \frac{-Low Good}{Total Low} * \log_2 \frac{Low Good}{Total Low} \right) + \frac{-Low Bad}{Total Bad} * \log_2 \left( \frac{Low Bad}{Total Bad} \right)$$

$$= \frac{1}{5} * \left( \frac{0}{1} * \log_2 * \frac{0}{1} \right) + \frac{-1}{1} * \log_2 \left( \frac{1}{1} \right)$$

$$= \frac{1}{5} * (0 * \log_2 0) + (-1 * \log_2 1)$$

$$= \frac{1}{5} * (0 + 0)$$

$$= 0$$

Nilai Gain Total / Hasil Prediksi

$$= Entropy Total - (Nilai Gain + Nilai Medium + Nilai Low)$$

$$= 0,72193 - 0$$

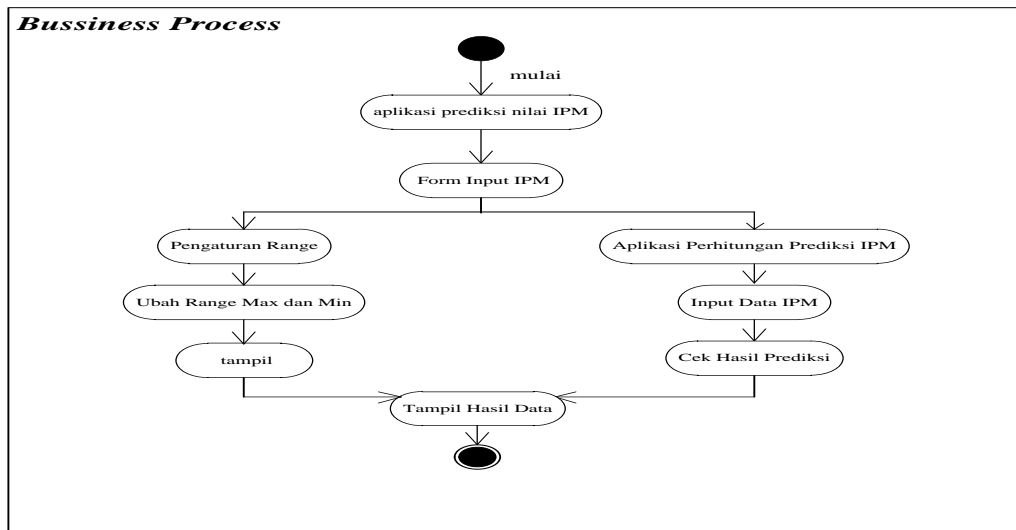
$$= 0,72193$$

### Perancangan dan Implementasi Sistem

Dalam perancangan sistem untuk Indeks Prediksi Manusia ini digunakan alat bantu perancangan Unified Modeling Language (UML). Bahasa Pemodelan Terpadu (UML) adalah bahasa pemodelan visual yang dominan dalam pengembangan perangkat lunak berorientasi objek[8] yang terbukti dapat mengirimkan informasi di antara pengguna, pengembang, perancang dan manajer secara efisien [9].

**a. Bussiness Process**

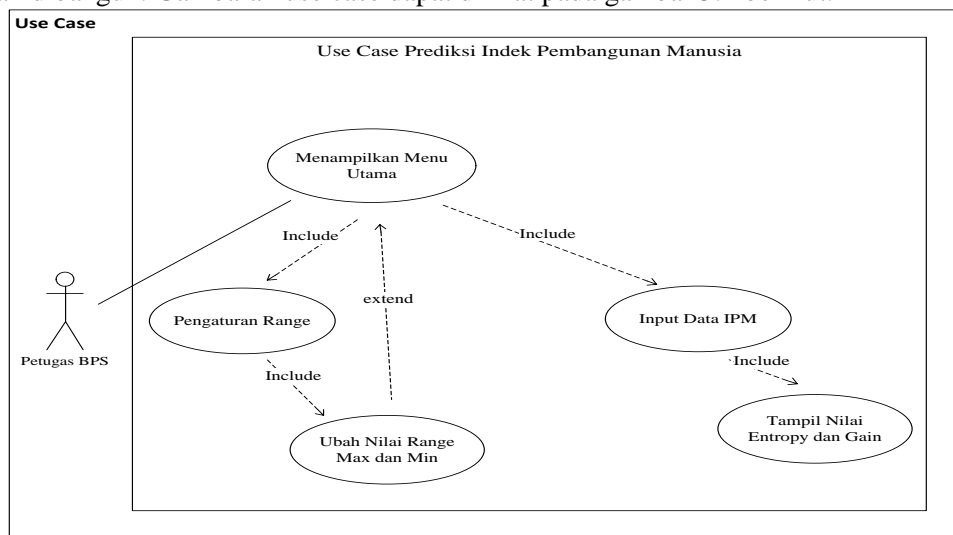
Kualitas model proses bisnis adalah sesuatu yang penting di bidang model berbasis pengembangan perangkat lunak[10]. Berikut ini penggambaran bussiness process yang diusulkan untuk prediksi nilai IPM, Dari Bussiness Process Model ini dapat dijelaskan bahwa hal yang pertama sistem lakukan adalah melakukan input data berupa harapan hidup, rata-rata lama sekolah, harapan berapa lama sekolah serta nilai pengeluaran perkapita. Melalui algoritma decision tree C4.5 nilai nilai tersebut akan di proses untuk mendapatkan nilai IPM . Berikut ini adalah gambaran business process untuk prediksi nilai IPM ini dapat dilihat pada gambar 5.1 berikut ini:



Gambar 3.1 Bussines Process Aplikasi Prediksi IPM

**b. Use Case Diagram**

Use case diagram digunakan untuk menggambarkan hubungan antara actor dan system yang akan dibangun. Gambaran use case dapat dilihat pada gambar 5.2 berikut.



Gambar 3.2 Use Case Prediksi Indeks Pembangunan Manusia

Dari gambaran use case diatas maka scenario untuk jalannya use case diagram dijelaskan pada tabel 2.7 berikut ini.

Tabel 2.5 Skenario Use Case BPS

<b>Judul Use Case</b>	Prediksi Indek Pembangunan (IPM) di BPS Kampar	
<b>Aktor</b>	BPS	
<b>Tujuan</b>	Input Data Berdasarkan Pertahun	
<b>Gambaran</b>	TU BPS input data IPM	
<b>Aksi Aktor</b>		<b>Respon Sistem</b>
.1. Input data IPM berupa angka harapan hidup, rata-rata lama sekolah, harapan lama sekolah, pengeluaran perkapital yang disesuaikan dan pengaturan range IPM		2. Hasil IPM, Gain dan Entrophy
<b>Alternative Course</b>		Tidak ada
<b>Exceptional Flow</b>		Tidak ada
<b>Pre-Conditions</b>		-BPS input data
<b>Post-Conditions</b>		-Hitung data sukses

**Implementasi Sistem**

Berikut gambaran program untuk Indek Predikasi Manusia sejak tahun 2010 sampai dengan tahun 2014



Gambar 3.3 Hasil 1



Form Perhitungan IPM

### Aplikasi Perhitungan Indeks Pembangunan Manusia

Angka Harapan Hidup

---

Angka Harapan Hidup

Harapan Lama Sekolah

Rata-Rata Lama Sekolah

Pengeluaran Per Kapita

Hitung IPM

---

Indeks Pendidikan

Indeks Pengeluaran Per Kapita

---

Indeks Pembangunan Manusia

Angka IPM dalam Ratusan

Pengaturan Range

Tahun	IPM	IPM Level	Kategori	Entropy	Gain
2010	0.11035	Low	Bad	0.81128	0.81128
2011	0.11498	Medium	Good		
2012	0.11629	High	Good		
2013	0.11746	High	Good		

Hitung

Gambar 3.4 Hasil 2

Form Perhitungan IPM

### Aplikasi Perhitungan Indeks Pembangunan Manusia

Angka Harapan Hidup

---

Angka Harapan Hidup

Harapan Lama Sekolah

Rata-Rata Lama Sekolah

Pengeluaran Per Kapita

Hitung IPM

---

Indeks Pendidikan

Indeks Pengeluaran Per Kapita

---

Indeks Pembangunan Manusia

Angka IPM dalam Ratusan

Pengaturan Range

Tahun	IPM	IPM Level	Kategori	Entropy	Gain
2010	0.11035	Low	Bad	0.72193	0.72193
2011	0.11498	Medium	Good		
2012	0.11629	High	Good		
2013	0.11746	High	Good		
2014	0.11812	High	Good		

Hitung

Gambar 3.5 Hasil 3

Tahun	IPM	IPM Level	Kategori	Entrophy	Gain
2010	0.11035	Low	Bad	0.9183	0.45915
2011	0.11498	Medium	Bad		
2012	0.11629	Medium	Good		
2013	0.11746	Medium	Good		
2014	0.11812	High	Good		
2015	0.12121	High	Good		

Gambar 3.5 Hasil 4

#### 4. Kesimpulan

Berdasarkan dari pengolahan data yang penelitian lakukan dari 5 tahun terakhir 2010 sampai 2014 bisa dikatakan nilai IPM bagus dan nilai kategori tiap tahun naik berdasarkan pengujian yang dilakukan, nilai entropy IPM 0,72193 dan gain 0,72193 dimana nilai entropy didapatkan dari IPM level dan kategori yang hasil nilai entropy untuk nilai penjumlahan nilai gain. Gain adalah nilai tertinggi dari pengolahan data IPM yang memiliki hasil untuk mendapatkan nilai prediksi. Penelitian ini hendaknya dapat lebih disempurnakan dengan menambahkan metoda metoda untuk mendapatkan nilai sehingga didapat optimasi untuk mendapatkan nilai yang lebih optimal. Prediksi IPM menggunakan data 5 tahun terakhir dari data yang didapat di BPS Kampar, disarankan melakukan penelitian 6 tahun terakhir atau lebih dari 6 tahun.

#### Daftar pustaka

- [1] A. Davies and G. Quinlivan, "A panel data analysis of the impact of trade on human development," *J. Socio. Econ.*, vol. 35, no. 5, pp. 868–876, 2006.
- [2] A. Pratama and M. Al-Shaikh, "Relation and Growth of Internet Penetration Rate with Human Development Level from 2000 to 2010," *Commun. IBIMA*, vol. 2012, pp. 1–8, 2012.
- [3] A. D. Fajaryanto and A. Alamsyah, "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 O," vol. 3, no. 3, pp. 2938–2946, 2016.
- [4] D. M. Putra, "Pemodelan Indeks Pembangunan Manusia (IPM) Provinsi Jawa Timur Dengan Menggunakan Metode Regresi Logistik Ridge," vol. 4, no. 2, p. 96, 2015.

- [5] B. N. Sari and P. Priati, "Identifikasi Keterkaitan Variabel dan Prediksi Indeks Pembangunan Manusia (IPM) Provinsi Jawa Barat Menggunakan Dynamic Bayesian Networks," *J. INFOTEL - Inform. Telekomun. Elektron.*, vol. 8, no. 2, p. 150, 2016.
  - [6] A. E. Pramadhani and T. Setiadi, "PENERAPAN DATA MINING UNTUK KLASIFIKASI PREDIKSI PENYAKIT ISPA ( Infeksi Saluran Pernapasan Akut ) DENGAN ALGORITMA DECISION TREE ( ID3 )," *J. Sarj. Tek. Inform. e-ISSN 2338-5197*, vol. 2, no. 1, pp. 831–839, 2014.
  - [7] B. HSSINA, A. MERBOUHA, H. EZZIKOURI, and M. ERRITALI, "A comparative study of decision tree ID3 and C4.5," *Int. J. Adv. Comput. Sci. Appl.*, vol. 4, no. 2, pp. 13–19, 2014.
  - [8] J. Zheng, Y. Feng, and Y. Zhao, "A unified modeling language-based design and application for a library management information system," *Cybern. Inf. Technol.*, vol. 14, no. SpecialIssue, pp. 129–144, 2014.
  - [9] K. Siau, "An Analysis of Unified Modeling Language (UML) Graphical Constructs Based on BWW Ontology," *J. Database Manag.*, vol. 21, no. 1, pp. i–viii, 2010.
  - [10] M. Sadowska, "An Approach to Assessing the Quality of Business Process Models Expressed in BPMN," *Inform. Softw. Eng. J.*, vol. 9, no. 1, pp. 57–77, 2015.
-