

PENERAPAN ARIMA ENSEMBLE DALAM PERAMALAN KEBUTUHAN MAKANAN POKOK DI KABUPATEN JEMBER

¹Herman Efvendi, ²Lutfi Ali Muharom, ³Hardian Oktavianto
Program Studi Teknik Informatika Fakultas Teknik
Universitas Muhammadiyah Jember
Herman.efvendi89@gmail.com

ABSTRAK

ARIMA merupakan model peramalan yang termasuk dalam kelompok peramalan linier. ARIMA dapat digunakan pada data yang tidak terdapat pola musiman maupun memiliki pola musiman. Ada beberapa langkah untuk mendapatkan nilai peramalan menggunakan metode ARIMA yang sesuai dengan metodologi *Box-Jenkins*. Langkah-langkah tersebut meliputi identifikasi, estimasi parameter, uji diagnosa, dan peramalan.

Pada tugas akhir ini akan dilakukan pengujian terhadap *data testing* memperlihatkan bahwa nilai model ARIMA *ensemble* pada tahun 2011 adalah sebesar 35505,92, dengan selisih terhadap data *testing* sebesar 521,79. Pada tahun 2012, nilai pada model *ensemble* adalah sebesar 35108,22 dengan selisih terhadap data *testing* sebesar 2777,62. Nilai ini lebih besar dari nilai tahun 2012 pada model AR(1) dan ARMA(1,1). Hasil yang diperoleh pada model *ensemble* pada tahun 2011 dan 2012 menjadi tidak lebih baik dari model ARMA(1,1) dan AR(1), karena interval data yang terlalu lebar pada dua model ARMA, yaitu AR(1), ARMA(1,1). Model terbaik terletak pada model AR (1).

Kata Kunci : R Studio , ARIMA, ARIMA *Ensemble*

PENERAPAN ARIMA ENSEMBLE DALAM PERAMALAN KEBUTUHAN MAKANAN POKOK DI KABUPATEN JEMBER

¹Herman Efvendi, ²Lutfi Ali Muharom, ³Hardian Oktavianto
Program Studi Teknik Informatika Fakultas Teknik
Universitas Muhammadiyah Jember
Herman.efvendi89@gmail.com

ABSTRAK

ARIMA forecasting model which is included in the group of linear forecasting. ARIMA can be used on data that there is no seasonal pattern and has a seasonal pattern. There are several steps to get value ARIMA forecasting method in accordance with the Box-Jenkins methodology. Such steps include the identification, parameter estimation, diagnostic tests, and forecasting.

In this final project will be tested against the data testing showed that the value of the ARIMA model ensemble in 2011 amounted to 35505.92, with the difference to the data of testing of 521.79. In 2012, the value on the model ensemble is equal to the difference of the data 35108.22 testing at 2777.62. This value is greater than the value in 2012 on the model AR (1) and ARMA (1,1). The results obtained on the model ensemble in 2011 and 2012 being no better than a model ARMA (1,1) and AR (1), because the data is too wide interval on two models of ARMA, the AR (1), ARMA (1, 1). The best model is the model of the AR (1).

Kata Kunci : R Studio , ARIMA, ARIMA *Ensemble*

I. PENDAHULUAN

1.1 Latar Belakang

Pemerintah Indonesia melalui Badan Pusat Statistik (BPS, dahulu Biro Pusat Statistik) menyelenggarakan sensus secara nasional. Setiap sepuluh tahun sekali BPS menyelenggarakan sensus nasional, contoh data penduduk Indonesia : 1971 = 119208229, 1980 = 147490298, 1990 = 179378946, 2000 = 206264595, 2010 = 237641326 , selain itu BPS juga melakukan pengumpulan data, menerbitkan publikasi statistik nasional maupun daerah, serta melakukan analisis data statistik yang digunakan dalam pengambilan keputusan (BPS, 2014). Sejak pola pembangunan yang diarahkan pada otonomi daerah, dibutuhkan perhitungan statistik sampai dengan skala kecamatan bahkan desa untuk membantu pemerintah daerah dalam upaya mendongkrak pembangunan daerah.

Perguruan Tinggi sebagai salah satu pusat riset yang didalamnya terhimpun sarana dan prasarana IPTEK yang cukup mutakhir dan tersedia cukup sumber daya manusia berkualitas tinggi. Salah satu hambatan didalam penelitian atau riset adalah masalah ketersediaan data, salah satu contoh adalah data sensus dan survei. Data sensus dan survei sulit diakses, dengan birokrasi berbelit akan menghambat penelitian yang dilakukan oleh perguruan tinggi.

R adalah salah satu dari program sumber terbuka yang dapat diunduh gratis di <http://cran.r-project.org/>, yang mana situs resminya ada di <http://www.r-project.org/>. R pertama kali diciptakan oleh Ross Ihaka dan Robert Gentleman (-nama R berasal dari huruf pertama nama depan kedua orang ini), mereka adalah Statistawan asal Selandia Baru, sedangkan kini R dikembangkan sebagai upaya kolaborasi pakar-pakar statistik dan komputer di seluruh dunia. R diluncurkan pertama kali tahun 1997, dan versi terakhir ketika tulisan ini dibuat adalah 3.1.0. Bagi sebagian pengguna yang terbiasa mengguna GUI dengan sistem menu, R juga menyediakan banyak GUI yang

berbasiskan sistem menu, antara lain *R Studio*, *Tinn-R*, *R Commander* dan banyak lagi yang lainnya, dan dapat diunduh gratis juga.

R Studio adalah salah satu GUI untuk R, salah satu keunggulan pada *R Studio* ini adalah dapat dijalankan pada *browser*. Sehingga dengan dijalankan diatas *browser*, maka pengguna tidak memerlukan lagi instalasi R, kecuali paket (*package*) pemrograman sesuai dengan kebutuhan pengguna.

1.2 Rumusan Masalah

Berdasarkan latar belakang di atas, maka permasalahan dalam penelitian ini adalah sebagai berikut:

1. Bagaimana meramalkan pola konsumsi makanan pokok di Kabupaten Jember ?
2. Bagaimana mengimplementasikan metode ARIMA ensemble pada kebutuhan makanan pokok di Kabupaten jember?

1.3 Batasan masalah

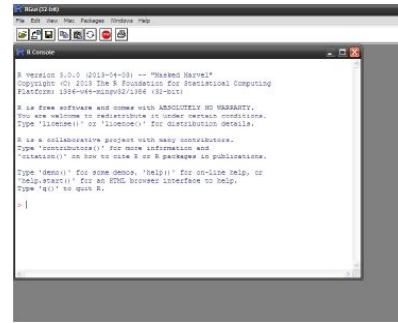
Untuk menghindari timbulnya salah penafsiran dan tidak menimbulkan penyimpangan dari masalah dan tujuan penelitian yang akan dilaksanakan, maka diperlukan adanya pembatasan masalah, yaitu:

1. Data yang dibangun adalah data yang bersumber dari survei BPS, yaitu: SUSENAS mulai tahun 1997 – 2012.
2. Data yang dipakai adalah data sensus Kabupaten Jember.
3. Uji coba implementasi menggunakan ARIMA *ensemble* untuk mengetahui pola konsumsi di kabupaten Jember.
4. Variabel yang digunakan dalam uji coba adalah pola konsumsi rumah tangga pada data makanan pokok yang ada pada SUSENAS.

1.4 Tujuan penelitian

Berdasarkan rumusan masalah di atas, maka tujuan dari penelitian ini adalah

1. Melakukan uji coba pada program *R Studio* dengan melakukan peramalan (*forecasting*) pada pola konsumsi pangan (pokok) di Kabupaten Jember.
2. Melakukan implementasi menggunakan *ARIMA ensemble* untuk analisa data.



Gambar 2.1 GUI Standar R

1.5 Manfaat Penelitian

Hasil penelitian ini diharapkan dapat bermanfaat untuk :

1. Sarana mempermudah belajar ilmu statistik, terutama tentang pemrograman *R Studio*.
2. Membangun pusat data penelitian yang memudahkan civitas akademika dalam melakukan penelitian/riset

II. TINJAUAN PUSTAKA

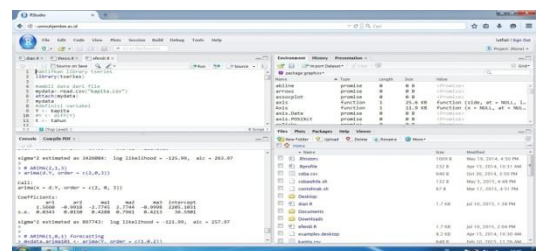
2.1 Program R

R adalah salah satu dari program sumber terbuka yang dapat diunduh gratis di <http://cran.r-project.org/>, yang mana situs resminya ada di <http://www.r-project.org/>. R pertama kali diciptakan oleh Ross Ihaka dan Robert Gentleman (nama R berasal dari huruf pertama nama depan kedua orang ini), mereka adalah statistikawan asal Selandia Baru, sedangkan kini R dikembangkan sebagai upaya kolaborasi pakar-pakar Statistikawan dan komputer di seluruh dunia. R diluncurkan pertama kali tahun 1997, dan versi terakhir saat ini adalah 3.1.2. Pada saat mengunduh R, untuk memilih nama-nama negara terdekat sebagai *server mirror* agar proses pengunduhan dapat lebih cepat, maka disarankan memilih Indonesia jika berada di Indonesia atau negara sekitarnya.

Selain gratis ada banyak kelebihan R yang lain yaitu :

1. R tersedia untuk berbagai sistem operasi selain *Windows*, seperti *Macintosh, Linux* dan *UNIX*
2. Memiliki kemampuan membuat grafis yang canggih
3. Sintaksnya mudah dipelajari dengan banyak fungsi-fungsi statistik yang terpasang
4. R dapat dengan mudah diperluas dengan menciptakan fungsi-fungsi buatan pengguna sendiri disamping juga tersedia banyak program-program kecil sebagai tambahan (*add in*), yang disebut paket (*package*) yang juga dapat diunduh gratis
5. R merupakan bahasa pemrograman komputer, sehingga bagi pemrogram menjadi lebih akrab, sedangkan bagi pemakai awal akan merupakan langkah yang mudah untuk memulai sebagai pemrogram komputer.

R Studio adalah salah satu GUI untuk R, salah satu keunggulan pada *R studio* ini adalah dapat dijalankan pada *browser*. Sehingga dengan dijalankan diatas *browser*, maka pengguna tidak memerlukan lagi instalasi R, kecuali paket (*package*) pemrograman sesuai dengan kebutuhan pengguna.



Gambar 2.2 R Studio Berbasis Web

Disamping kelebihan yang disebutkan di atas, R juga memiliki kekurangan yaitu:

1. R tidak menyediakan dukungan komersial
2. Karena R merupakan bahasa pemrograman maka diperlukan waktu membiasakan dengan sintaks-sintaksnya

2.2 Peramalan

Peramalan adalah penggunaan data masa lalu dari sebuah variabel atau kumpulan variabel untuk mengestimasi nilai di masa yang akan datang. Peramalan merupakan bagian vital bagi setiap organisasi dan untuk setiap pengambilan keputusan manajemen yang sangat signifikan. Peramalan menjadi dasar bagi perencanaan jangka panjang perusahaan.

2.2.1 Model Autoregressive Integrated Moving Average (ARIMA)

Pada model ARIMA (p,d,q) , nilai yang akan datang dari suatu variabel diasumsikan sebagai fungsi linier dari beberapa pengamatan di masa lalu dan *random error* (Khashei *et al.*, 2009). Model ARIMA (p,d,q) secara umum, yaitu (Wei, 2009) :

$$\Phi_p(B)(1 - B)^d Z_t = \theta_0 + \theta_q(B) a_t \quad (2.1)$$

Dengan

$$\Phi_p(B) = (1 - \phi_1 B - \dots - \phi_p B^p), \theta_q(B) = (1 - \theta_1 B - \dots - \theta_q B^q)$$

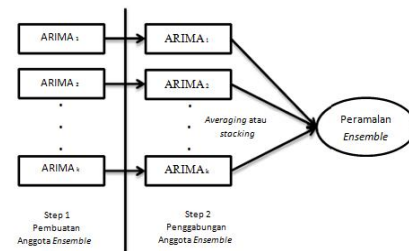
dan θ_0 adalah *intercept* pada model untuk *difference* ke- d

Pembuatan model ARIMA dapat dilakukan dengan menggunakan tiga prosedur yaitu identifikasi model, estimasi model dan cek diagnosa (Jenkins, 2006). Identifikasi model merupakan metodologi untuk

mengidentifikasi perlunya suatu transformasi seperti transformasi untuk stasioner dalam varians, *transformasi differencing*, keputusan untuk memasukkan parameter θ_0 ketika $d > 0$ dan penentuan order p dan q pada ARIMA (Wei, 2009).

2.2.2 Model Autoregressive Integrated Moving Average Ensemble (ARIMA Ensemble)

ARIMA *ensemble* merupakan penggabungan hasil ramalan beberapa model ARIMA. Pembentukan ARIMA *ensemble* terdiri dari dua langkah. Pertama, menciptakan anggota *ensemble* dari beberapa model ARIMA



Gambar 2.3 Arsitektur ARIMA Ensemble

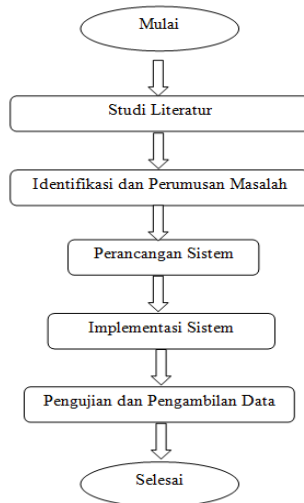
selanjutnya menggabungkan hasil ramalan anggota *ensemble* dari ARIMA yang terbentuk dengan menggunakan *averaging* dan *stacking* sehingga didapatkan hasil ramalan ARIMA *ensemble*. Arsitektur model ARIMA *ensemble* dapat dilihat melalui Gambar 2.3

Beberapa pendekatan telah diperkenalkan untuk membangkitkan anggota- anggota *ensemble* dan secara umum dapat terbagi dalam dua kelompok utama, yaitu:

1. Membuat suatu himpunan model dari berbagai variasi model peramalan dengan menjaga data *training* tidak berubah. Kelompok ini dikenal dengan randomisasi (Shu dan Burn, 2004)
2. Merubah suatu himpunan data *training*. Kelompok kedua ini

biasanya dilakukan dengan menggunakan teknik-teknik *resampling* (Breiman, 1996).

III. METODOLOGI PENELITIAN



3.1 Sumber Data

Data yang digunakan pada penelitian ini merupakan data yang didata oleh Badan Pusat Statistik (BPS) yang diambil dari website <http://ada.edu>. Terdapat beberapa data sensus, untuk penelitian ini data yang diambil adalah data SUSENAS mulai tahun 1997-2012. Data tersebut diambil setiap tahun oleh BPS dan mempunyai variabel yang hampir sama setiap tahunnya (lihat pada Lampiran 1).

3.2 Variabel Penelitian

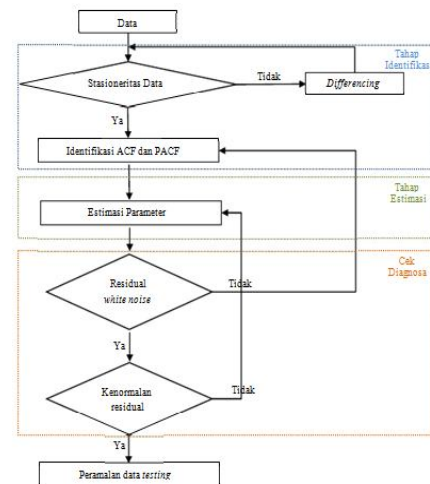
Variabel penelitian yang akan digunakan pada penelitian ini adalah data belanja atau konsumsi kebutuhan pokok (beras, jagung) masyarakat pada Kabupaten Jember, Propinsi Jawa Timur per minggu. Data sensus per tahun di modelkan dengan rata-rata (*averaging*), diambil dari data SUSENAS mulai tahun 1997 sampai dengan tahun 2012.

Data rata-rata belanja kebutuhan pokok dibagi menjadi dua yaitu data *training* dan data *testing*. Data *training* digunakan untuk memodelkan yaitu mulai tahun 1997 sampai dengan tahun 2010. Sedangkan data *testing* (tahun 2011 – tahun 2012) digunakan untuk

membandingkan data yang dihasilkan dari model ARIMA dan ARIMA *ensemble*.

3.3 Langkah Penelitian

Pembuatan model ARIMA dapat dilakukan dengan menggunakan tiga prosedur yaitu identifikasi model, estimasi dan uji diagnostik. Langkah – langkah dalam model ARIMA ini dapat digambarkan sebagai berikut.



Gambar 3.2 Langkah-langkah Model ARIMA

Berikut ini merupakan penjelasan dari setiap langkah-langkah yang digambarkan pada Gambar 3.2:

1. Membagi data menjadi dua bagian, yaitu data *training* dan *testing*. Data *training* digunakan untuk memodelkan, sedangkan data *testing* digunakan untuk membandingkan nilai dari setiap model.
2. Memeriksa stasioneritas data dalam varian melalui plot *time series* (tidak stasioner jika terdapat pola corong). Jika data tidak stasioner maka dilakukan proses *differencing*. Identifikasi ACF dan PACF yang akan menentukan nilai orde p , q , P , Q dan S pada model ARIMA.
3. Pada tahap uji diagnostik yang dilakukan meliputi pengujian *white noise* dari ϵ_t (residual) dan uji distribusi normal dari ϵ_t . Jika residual tidak *white noise*, maka dilakukan identifikasi ulang untuk penentuan orde pada model ARIMA. Sedangkan

jika residual tidak berdistribusi normal, maka dilakukan penanganan *outlier*.

4. Meramalkan data *testing*.

IV. HASIL DAN PEMBAHASAN

4.1 Uji Program

Pada penelitian ini digunakan data pada variabel *household expenditure* (belanja rumah tangga) mulai tahun 1997 sampai dengan 2012. Pada variabel *household expenditure*, data yang digunakan adalah data konsumsi pangan yang meliputi : makanan pokok (beras/jagung/tepung). Data tersebut dapat dilihat pada tabel dibawah ini.

Tabel 4.1 Data SUSENAS Kabupaten Jember mengenai *Household Expenditure*

No	Tahun	Makanan pokok	Makanan Instan
1	1997	6701,80	656,33
2	1998	11392,59	891,94
3	1999	13879,86	1041,22
4	2000	16083,39	1127,55
5	2001	16367,13	1190,51
6	2002	16609,00	1321,99
7	2003	17250,87	1853,47
8	2004	17450,87	2053,47
9	2005	19550,13	1829,91
10	2006	28258,13	2102,59
11	2007	30584,18	3097,75
12	2009	33440,44	3415,58
13	2010	35643,96	3501,90
14	2011	36027,71	3664,86
15	2012	37885,84	5633,38

4.2 Model ARIMA Ensemble

Model ARIMA *ensemble* merupakan model *ensemble* yang anggotanya terdiri dari beberapa model ARIMA. Terdapat dua kombinasi pada *ensemble* yaitu *averaging* dan *stacking*. Pada penelitian ini, dilakukan *ensemble averaging* pada semua data dan hasil *training*. Pada pembahasan model ARIMA terdapat 2 model ARIMA yang dijadikan sebagai model didalam *ensemble*, hasil dari semua data *training* yang dilakukan pada model ARMA kemudian akan diambil rata-rata (*average*).

Nilai pada ARIMA *ensemble* diperoleh dari rata-rata pada tiga model ARMA, yaitu : AR(1) dan ARMA(1,1).

Maka diperoleh nilai ARIMA *ensemble* pada tahun 2011 dan tahun 2012 adalah sebagai berikut :

Tabel 4.8 Hasil ARIMA Ensemble

Model	Tahun	Hasil	Testing	Selisih (Testing - Hasil)
AR(1)	2011	35643,93	36027,71	383,78
	2012	35643,89	37885,84	2241,95
ARMA(1,1)	2011	35367,90	36027,71	659,81
	2012	34572,54	37885,84	3313,30
ARIMA ensemble	2011	35505,92	36027,71	521,79
	2012	35108,22	37885,84	2777,62

Model	Pengujian Asumsi White Noise			Pengujian Asumsi Normal	
	Lag	p-value	Ket	p-value	Ket
	1	< 2x10 ⁻¹⁶	Signifikan		
	2	0,01009	Signifikan		
MA(2)	0	NA	Tidak Signifikan	< 2x10 ⁻¹⁶	Tidak Signifikan
	1	< 2x10 ⁻¹⁶	Tidak Signifikan		
	2	0,08781	Signifikan		
	0	NA	Tidak Signifikan		
AR(2,2)	1	NA	Tidak Signifikan	NA	Tidak Signifikan
	2	NA	Signifikan		
AR(2,1)	0	NA	Tidak Signifikan	0,695	Tidak Signifikan
	1	< 2x10 ⁻¹⁶	Tidak Signifikan		
	2	0,9178	Signifikan		
	0	NA	Tidak Signifikan		
AR(2)	1	< 2x10 ⁻¹⁶	Tidak Signifikan	0,552	Tidak Signifikan
	2	0,1753	Signifikan		

4.3 Peramalan

Berdasarkan data pada tabel 4.5 pada uji diagnosa, model ARMA yang memungkinkan AR(1), ARMA(1,1), maka uji peramalan akan dilakukan dengan meramalkan data dua tahun kedepan, hasilnya sebagai berikut :

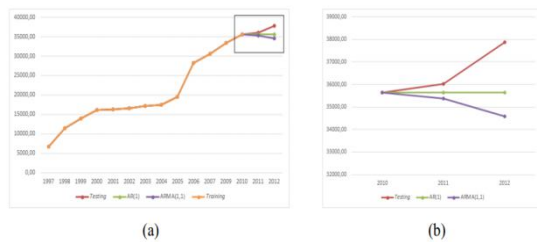
Tabel 4.6 Hasil Peramalan Model

MODEL	Tahun ke 1	Tahun ke 2
AR (1)	35643,93	35643,89
ARMA (1,1)	35367,90	34572,54

Dari data tersebut kemudian dibandingkan dengan data *testing* (Tahun 2011 dan 2012), maka akan diperoleh hasil sebagai berikut.

Model	Tahun	Hasil	Testing	Selisih (Testing - Hasil)
AR(1)	2011	35643,93	36027,71	383,78
	2012	35643,89	37885,84	2241,95
ARMA(1,1)	2011	35367,90	36027,71	659,81
	2012	34572,54	37885,84	3313,30

Grafik perbandingan Uji *Testing*.



V. PENUTUP

5.1 Kesimpulan

Berdasarkan pembahasan dan permasalahan pada penelitian ini terdapat beberapa kesimpulan, yaitu:

1. Uji coba peramalan pada pola konsumsi (makanan pokok) di Kabupaten Jember, diperoleh model terbaik peramalan menggunakan model AR (1). Keterbatasan data bisa menjadi menyebabkan *ensemble* bukan yang terbaik.
2. Dari hasil perbandingan dengan *data testing* terlihat bahwa pada model AR(1) mempunyai selisih terkecil pada tahun 2011 yaitu sebesar 383, sedangkan peramalan pada tahun 2012, data yang mempunyai selisih terkecil adalah model AR(1) yaitu sebesar 2241,95 di Kabupaten Jember.

5.2 Saran

Saran yang dapat disampaikan untuk pengembangan dari penelitian ini adalah mencoba simulasi menggunakan model yang lain misal *Exponential Smoothing*, *Neural Network*.

DAFTAR PUSTAKA

- Bishop, C.M. 1995. *Neural Networks for Pattern Recognition*. New York: Oxford
- Box, G.E.P., Jenkins, G.M. 1976. *Time series Analysis Forecasting and Control Revised Edition*. Oakland: Holden-Day, Inc.
- Breiman, L. 1996. *Stacked regression*.

Machine Learning, Vol. 24: 59-64.

Biro Pusat Statistik, <http://www.bps.go.id>. [25 - 28 Maret 2014].

FAO. 2012. *"The State of Food Insecurity in The World 2012."* Rome: FAO.

Inmon, W.H., 2005. *Building Data warehouse*, 4th Edition. John Wiley & Sons, Canada.

J. Supranto, 2000, *Statistik Teori dan Aplikasi*, Erlangga, Jakarta.

Kementerian Pertanian. 2012. Laporan Kinerja Kementerian Pertanian Tahun 2011. http://www.pertanian.go.id/sakip/admin/data2/LAKIP_KEMANTAN_211.pdf. [20 Maret 2014].

Khashei, M., Bijari, M., Ardali, G.A.R. 2007. *Improvement of Auto-Regressive Integrated Moving average Models Using Fuzzy Logic and Artificial Neural Networks*. Neurocomputing. Vol. 72 : 4-6.

Lane, Paul., 2002. *Oracle9i Data Warehousing Guide*. Release 2. Oracle Corporation

Mukhlis F., Bayu Adhi T., 2011. Rancang Bangun *Data Warehouse* Untuk Menunjang Evaluasi Akademik Di Fakultas. Prosiding Konferensi Nasional Teknologi Informasi dan Aplikasinya (KNTIA).

Prayoto,. 2008. "Peranan Perguruan Tinggi Dalam Pengembangan IPTEK." Tidak Diterbitkan. Makalah. Yogyakarta: Seminar Nasional Dies Natalis 45 UGM