

## IMPLEMENTASI ALGORITMA REGRESI LINEAR SEDERHANA DALAM MEMPREDIKSI BESARAN PENDAPATAN DAERAH (STUDI KASUS: DINAS PENDAPATAN KAB. DELI SERDANG)

Fransiskus Ginting<sup>1</sup>, Efori Buulolo<sup>2</sup>, Edward R Siagian<sup>3</sup>

Program Studi Teknik Informatika STMIK Budi Darma, Medan, Indonesia

Email: Aanginting1512@gmail.com,

### Abstrak

Data Mining merupakan suatu penemuan informasi dengan melakukan penggalian pola informasi yang berisi pencarian *trend* dalam sejumlah data yang sangat besar serta membantu proses penyimpanan data dalam mengambil suatu keputusan di waktu yang akan datang. Dalam menentukan pola dilakukan teknik klasifikasi mengumpulkan *record* (Training set). Pendapatan daerah umumnya berasal dari pajak dan retribusi daerah, pajak daerah merupakan salah satu sumber pendanaan bagi daerah secara rata-rata nasional belum mampu memberikan kontribusinya yang besar bagi pembentukan pendapatan asli daerah. Dengan memanfaatkan data Pendapatan Daerah dapat menghasilkan peramalan dan prediksi penghasilan Pendapatan Daerah kedepannya agar sesuai dengan kenyataan/realitas sehingga RAPBD yang telah direncanakan dapat berjalan dengan lancar. Regresi Linear Sederhana atau sering disingkat dengan SLR (Simple Linear Regression) merupakan salah satu metode statistik yang dipergunakan dalam produksi untuk melakukan peramalan ataupun prediksi tentang karakteristik kualitas maupun kuantitas untuk menggambarkan proses yang terkait dengan pengolahan data perolehan besaran pendapatan daerah. Sehingga dalam tahap pengujian dengan *visual basic net* dapat membantu dalam mengolah data Besaran Pendapatan Daerah yang valid.

**Kata Kunci:** Data Mining, Pendapatan Daerah, Algoritma Regresi Linear Sederhana, Visual Basic net 2008

### Abstract

Data Mining is an information discovery by extracting information patterns that contain trend searches in a very large amount of data and assist the process of storing data in making a decision in the future. In determining the pattern classification techniques do to collect records (Training set). Regional income is generally derived from local taxes and levies, local taxes are one source of funding for the region on the national average has not been able to make a large contribution to the formation of local revenue. By utilizing Regional Revenue data, it can produce forecasting and predictions of Regional Revenue income in the future to match the reality / reality so that the planned RAPBD can run smoothly. Simple Linear Regression or often abbreviated as SLR (Simple Linear Regression) is one of the statistical methods used in production to make predictions or predictions about the characteristics of quality and quantity to describe the processes associated with data processing for the acquisition of regional income. So that in the testing phase with visual basic net can help in processing valid Regional Revenue Amount data.

**Keywords:** Data Mining, Local Revenue, Simple Linear Regression Algorithm, Visual Basic net 2008

## 1. PENDAHULUAN

Prediksi adalah suatu proses memperkirakan secara sistematis tentang sesuatu yang paling mungkin terjadi di masa depan berdasarkan informasi masa lalu dan sekarang yang dimiliki, agar kesalahannya (selisih antara sesuatu yang terjadi dengan hasil perkiraan) dapat diperkecil.

Dinas Pendapatan Daerah adalah Dinas yang bertanggung jawab penuh atas pemungutan jasa atau perizinan tertentu yang khusus disediakan dan diberikan oleh pemerintah daerah untuk kepentingan pribadi dan badan. Dinas Pendapatan Daerah memiliki berbagai tugas pokok, yang terdiri atas: melaksanakan urusan administrasi yang meliputi pengelolaan surat menyurat, kepegawaian, keuangan, perlengkapan dan kerumahtanggaan serta mengkoordinasikan penyusunan program dan penyelenggaraan tugas-tugas bidang secara terpadu.

Beberapa faktor-faktor yang mempengaruhi besaran pendapatan daerah yaitu pengeluaran pemerintah, jumlah penduduk, dan inflasi. Sehingga setiap akhir tahun, dinas pendapatan Deli Serdang melakukan rapat dengan tujuannya adalah untuk menentukan sumber pendapatan, sumber potensi pendapatan yang baru serta besaran pendapatan. Hanya saja cara tersebut sering tidak sesuai dengan kenyataan/realitas, akibatnya berpengaruh pada belanja daerah yang sudah ditetapkan sebelumnya, sehingga kadang ada program pemerintah daerah yang dijalankan tidak sesuai dengan yang telah direncanakan.

Memanfaatkan data mining dan Algoritma Regresi Linear Sederhana dengan menggunakan data yang lama dan besar untuk mengetahui besaran pendapatan Kabupaten Deli Serdang. Algoritma Regresi Linear Sederhana suatu metode statistik yang berfungsi untuk menguji sejauh mana hubungan sebab akibat antara variabel faktor penyebab (X) terhadap variabel akibatnya. Faktor penyebab pada umumnya dilambangkan dengan X atau disebut juga dengan *predictor* sedangkan variabel akibat dilambangkan dengan Y atau disebut juga dengan *response*. Regresi Linear Sederhana atau sering disingkat dengan SLR (Simple Linear Regression) juga merupakan salah satu metode statistik yang dipergunakan dalam produksi untuk melakukan peramalan ataupun prediksi tentang karakteristik kualitas maupun kuantitas.

Pada penelitian terdahulu oleh Murni Marbun, Hengki Tamando Sihotang, Melda Agustina Nababan yang berjudul Perancangan Sistem Peramalan Jumlah Wisata Asing (Studi Kasus: Badan Pusat Statistik jumlah wisatawan asing di Sumatera) di Fakultas STMIK Pelita Nusantara mengatakan "Metode Regresi Linear berhasil diterapkan untuk

peramalan jumlah wisatawan asing di Sumatera Utara berdasarkan data historis 10 tahun terakhir menunjukkan bahwa wisatawan asing yang akan datang berkunjung pada Januari tahun 2016 di Sumatera Utara adalah 16.937 jiwa” [1].

Berdasarkan penelitian oleh Anwar, Dosen Program Studi Pendidikan Matematika FKIP Unsyiah yang berjudul Mendapatkan Algoritma Menghitung Koefisien-koefisien Persamaan Regresi Linear Dua Variabel untuk Dikoding ke dalam Bahasa Pemrograman mengatakan “Perhitungan koefisien a dan b dari persamaan regresi  $y = a + bx$  secara sangat sederhana sekali dapat dilakukan dengan membuat program sendiri oleh para pembaca. Dengan cara ini kita akan memiliki program yang dapat dimodifikasikan sendiri tanpa harus terkendala oleh program yang telah dibuat oleh orang lain” [2].

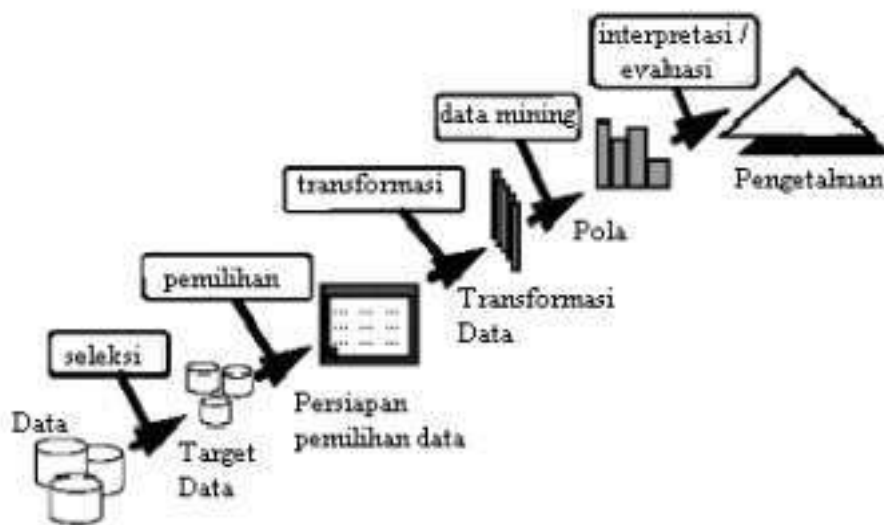
## 2. LANDASAN TEORI

### 2.1 Data Mining

Data mining adalah suatu istilah yang digunakan untuk menguraikan penemuan pengetahuan di dalam database. Data mining adalah proses yang menggunakan teknik statistik, matematika, kecerdasan buatan, dan *machine learning* untuk mengekstraksi dan mengidentifikasi informasi yang bermanfaat dan pengetahuan yang terakut dari berbagai database besar [3].

Data mining juga dikenal dengan nama *Knowledge Discovery In Database* (KDD). Selain itu Data mining juga didefinisikan sebagai suatu set teknik yang dipergunakan secara otomatis dalam mengeksplorasi keseluruhan dan membawa kepermukaan relasi-relasi yang kompleks pada set data yang sangat besar. Set data yang dimaksud disini adalah set data yang berbentuk tabulasi, seperti yang banyak diimplementasikan dalam teknologi manajemen basis data relasional.

Proses data mining atau yang biasa disebut dengan *Knowledge Discovery Databases* (KDD). KDD sering kali digunakan secara bergantian untuk menjabarkan proses penggalian suatu informasi tersembunyi dalam suatu basis data yang besar. Sebenarnya kedua istilah tersebut mempunyai konsep yang berbeda, akan tetapi berkaitan satu dengan yang lain. Dan satu tahapan dalam keseluruhan proses KDD seperti terlihat pada gambar 3.2 di bawah ini:



Gambar 1. Proses Knowledge Discovery Database (KDD)

Sumber: Dicky Nofriansyah, 2014 [4]

### 2.2 Algoritma Regresi Linear Sederhana

Regresi linear sederhana adalah analisis regresi yang melibatkan hubungan antara satu variabel tak bebas dihubungkan dengan satu variabel bebas. Regresi linier juga merupakan metode statistik yang berfungsi untuk menguji sejauh mana hubungan sebab-akibat antara variabel faktor penyebab (x) terhadap variabel akibatnya. Faktor penyebab pada umumnya dilambangkan dengan X sedangkan variabel akibat dilambangkan dengan Y. Regresi linear sederhana atau sering disingkat dengan SLR (Simple Linier Regression) juga merupakan salah satu metode statistik yang dipergunakan dalam produksi untuk melakukan peramalan atau pun prediksi tentang karakteristik kualitas maupun kuantitas. Persamaan umum metode regresi linier sederhana dalam penelitian ini adalah:  $Y = a + b(X)$  Keterangan: a = Konstanta b = Koefisien regresi Y = Variabel dependen (variabel tak bebas) X = Variabel independen (variabel bebas).

Menentukan koefisien persamaan a dan b dapat dengan menggunakan metode kuadrat terkecil, yaitu cara yang dipakai untuk menentukan koefisien persamaan dan dari jumlah pangkat dua (kuadrat) antara titik-titik dengan garis regresi yang dicari yang terkecil. Dengan demikian, dapat ditentukan (Murni Marbun, 2018):  
Menghitung konstanta:

$$a = \frac{(\sum Y)(\sum X^2) - (\sum X)(\sum XY)}{n \sum X^2 - (\sum X)^2} \quad (1)$$

Menghitung koefisien:

$$b = \frac{n(\sum XY) - (\sum X)(\sum Y)}{n \sum X^2 - (\sum X)^2} \quad (2)$$

Langkah-langkah dalam melakukan Analisis Regresi Linear Sederhana:

1. Menentukan tujuan dari melakukan analisis Regresi Linear yaitu mempelajari hubungan yang diperoleh dan dinyatakan dalam persamaan matematika yang menyatakan hubungan antar variabel.
2. Mengidentifikasi Variabel Faktor Penyebab (X) dan Variabel Akibat (Y)
  - a. Variabel Faktor Penyebab (X) : Jumlah Periode
  - b. Variabel Faktor Akibat (Y) : Jumlah Besar Pendapatan Daerah
3. Melakukan Pengumpulan Data.

### 3. ANALISA DAN PERANCANGAN

#### 3.1 Analisa

Analisa dapat didefinisikan sebagai penguraian dari suatu sistem informasi yang utuh kedalam bagian-bagian komponennya dengan maksud untuk mengidentifikasi dan mengevaluasi permasalahan, hambatan, yang terjadi dalam kebutuhan yang diharapkan sehingga dapat diusulkan perbaikan-perbaikannya. Penelitian dilakukan di Dinas Pendapatan Daerah Kabupaten Deli Serdang, dan berdasarkan penelitian yang dilakukan pengolahan data untuk memperediksi besaran pendapatan daerah agar sejalan dengan rancangan anggaran pendapatan dan belanja daerah.

Analisa dalam memprediksi besaran pendapatan daerah di Dinas Pendapatan Daerah Kabupaten Deli Serdang sangat perlu dilakukan. Analisa ini berguna untuk mendapatkan pokok-pokok permasalahan yang sebenarnya menjadi inti permasalahan yang sering terjadi di Dinas Pendapatan Daerah Kabupaten Deli Serdang. Dalam hal ini perlu melakukan analisa data besaran pendapatan daerah kabupaten deli serdang pertahun, dimana data yang digunakan adalah data besaran pendapatan daerah kabupaten deli serdang tahun 2017 dan 2018 yang tercantum dalam laporan pertanggung jawaban bendahara penerima (SPJ Pendapatan-Fungsional) di Dinas Pendapatan Daerah Kabupaten Deli Serdang.

Analisa ini bertujuan agar Dinas Pendapatan Daerah Kabupaten Deli Serdang dapat mengetahui langkah-langkah yang harus dilakukan dalam pendataan sumber-sumber pendapatan daerah dan potensi pendapatan daerah baru, sehingga program pemerintahan yang telah direncanakan dapat berjalan dengan lancar sesuai dengan rancangan anggaran pendapatan belanja daerah kabupaten deli serdang.

Data yang digunakan untuk mengestimasi data besaran pendapatan daerah Kabupaten Deli Serdang adalah data tahun 2017 dan 2018, yang akan dibagi menjadi 2 bulan, Triwulan(3 bulan), Caturwulan(4 bulan), 1 semester(6 bulan), dan 1 tahun(12 bulan) untuk memudahkan analisa berikutnya.

Berikut adalah tabel besaran pendapatan daerah kabupaten Deli Serdang dapat dilihat pada tabel berikut ini:

Tabel 1. Data Besaran Pendapatan

No	Bulan(X)	Jumlah Pendapatan/dalam juta(Y)
1	2 bulan 2017	42.656.854.914,15
2	3bulan 2017	70.612.092.375,33
3	4bulan 2017	103.216.361.575,54
4	6bulan 2017	173.238.202.000,42
5	12/ 1tahun 2017	557.559.474.513,97
6	2bulan 2018	60.267.298.225,02
7	3bulan 2018	90.687.788.773,83
8	4bulan 2018	129.561.820.104,57
9	6bulan 2018	198.609.703.609,81
10	12 / tahun 2018	562.657.701.076,63

Di bawah ini merupakan tabel penyerdehanaan dari data besaran pendapatan kabupaten deli serdang.

Tabel 2. Penyerdehanaan Data Besaran Pendapatan

No	X	Y
1	2	42.656,80
2	3	70.612,10
3	4	103.216,36
4	6	173.238,20
5	12	557.559,47
6	2	60.267,30
7	3	90.687,70
8	4	129.561,82

9	6	198.609,70
10	12	562.657,70

Di bawah ini merupakan tabel nilai perhitungan untuk mencari nilai  $X^2$ ,  $Y^2$ , dan  $XY$  serta total dari masing-masing variabel

Tabel 3. Nilai Perhitungan  $X^2$ ,  $Y^2$ ,  $XY$ 

No	X	Y	$X^2$	$Y^2$	XY
1	2	42.656,80	4	1.819.602.586,24	85.313,6
2	3	70.612,10	9	4.986.068.666,41	211.836,3
3	4	103.216,36	16	10.653.616.971,65	412.865,44
4	6	173.238,20	36	30.011.473.939,24	1.039.429,2
5	12	557.559,47	144	310.872.562.586,68	6.690.713,64
6	2	60.267,30	4	3.632.147.449,29	120.534,6
7	3	90.687,70	9	8.224.258.931,29	272.063,1
8	4	129.561,82	16	16.786.265.201,71	518.247,28
9	6	198.609,70	36	39.445.812.934,09	1.191.658,2
10	12	562.657,70	144	316.583.687.369,29	6.751.892,4
Total( $\Sigma$ )	54	1.989.067,15	418	743.015.496.635,89	17.294.553,76

Maka langkah selanjutnya adalah menghitung nilai konstanta (a) dan koefisien (b) berdasarkan rumus di atas.

Menghitung konstanta a

$$a = \frac{(\Sigma Y)(\Sigma X^2) - (\Sigma X)(\Sigma XY)}{n \Sigma X^2 - (\Sigma X)^2}$$

$$a = \frac{(1.989.067,15)(418) - (54)(17.294.553,76)}{10(418) - (54)^2}$$

$$a = \frac{831.430.068,7 - 933.905.903,04}{4180 - 2916}$$

$$a = \frac{102.475.834,34}{1264}$$

$$a = 81.072,65$$

Menghitung Koefisien b

$$b = \frac{n(\Sigma XY) - (\Sigma X)(\Sigma Y)}{n \Sigma X^2 - (\Sigma X)^2}$$

$$b = \frac{10(17.294.553,76) - (54)(1.989.067,15)}{10(418) - (54)^2}$$

$$b = \frac{172.945.537,6 - 107.409.626,1}{4180 - 2916}$$

$$b = \frac{65.535.911,5}{1264}$$

$$b = 51.848,03$$

Maka persamaan regresinya adalah:

$$Y = a + bX$$

$$Y = 81.072,65 + 51.848,03X$$

Jadi, Peramalan nilai *estimasi* untuk besaran pendapatan pertahun kedepannya adalah:

$$Y = 81.072,65 + 51.848,03X$$

$$Y = 81.072,65 + 51.848,03(12)$$

$$Y = 81.072,65 + 622.176,36$$

$$Y = 703.249,01$$

Maka, nilai  $Y = 703.249,01$  yang di atas merupakan nilai *estimasi* besaran pendapatan daerah kabupaten deli serdang pada tahun 2019 dalam nilai asli sekitar Rp. 703.249.000.000,01

#### 4. IMPLEMENTASI

*Input* data merupakan tahapan pengisian data sehingga data yang dikehendaki dapat diterima oleh sistem. Tampilan *input* data form pendapatan dapat dilihat dari gambar berikut ini :

Node	Bulan	Tahun	Pendapatan	Kurang
K001	1 Bulan	2017	1000000000	1000000000
K002	2 Bulan	2017	2000000000	2000000000
K003	3 Bulan	2017	3000000000	3000000000
K004	4 Bulan	2017	4000000000	4000000000
K005	5 Bulan	2017	5000000000	5000000000
K006	6 Bulan	2017	6000000000	6000000000
K007	7 Bulan	2017	7000000000	7000000000
K008	8 Bulan	2017	8000000000	8000000000
K009	9 Bulan	2017	9000000000	9000000000
K010	10 Bulan	2017	10000000000	10000000000

Gambar 2. Tampilan Form Pendapatan

Keterangan :

Gambar 5.2 adalah data yang akan dimasukkan dalam proses data mining, yaitu hasil transformasi data 2017 dan 2018 di Dinas Pendapatan Daerah Kabupaten Deli Serdang. Sebelum menentukan node terpilih, maka terlebih dahulu dilakukan *input* data jumlah besaran pendapatan yang diperoleh dari *database* besaran pendapatan daerah kedalam *Microsoft Excel* dan diinputkan ke dalam *visual basic* 2008 untuk selanjutnya dilakukan proses data.

#### 1. Tampilan Proses

Menu proses merupakan hasil dari nilai input data, dapat di lihat pada gambar di bawah ini:

No	X	Y	X2	XY
1	2	4000000000	4	16000000000000
2	3	7000000000	9	21000000000000
3	4	10000000000	16	28000000000000
4	5	13000000000	25	32500000000000
5	6	16000000000	36	38400000000000
6	7	19000000000	49	44100000000000
7	8	22000000000	64	50400000000000
8	9	25000000000	81	57600000000000
9	10	28000000000	100	65600000000000
10	11	31000000000	121	74610000000000
Total	54	180000000000	410	74610000000000

Gambar 3. Tampilan Menu Proses

Keterangan:

Gambar 3. adalah menu proses dari tahap pencarian hasil nilai koefisien variabel-variabel yang diperlukan dalam proses perhitungan pada form hasil prediksi.

#### 2. Tampilan Output

*Output* data merupakan tahapan hasil yang dicapai dari sistem yang menghasilkan suatu nilai estimasi. Tampilan *output* data form hasil prediksi dapat dilihat dari gambar 3. berikut ini :

No	X	Y	X2	XY
1	2	4000000000	4	16000000000000
2	3	7000000000	9	21000000000000
3	4	10000000000	16	28000000000000
4	5	13000000000	25	32500000000000
5	6	16000000000	36	38400000000000
6	7	19000000000	49	44100000000000
7	8	22000000000	64	50400000000000
8	9	25000000000	81	57600000000000
9	10	28000000000	100	65600000000000
10	11	31000000000	121	74610000000000
Total	54	180000000000	410	74610000000000

Gambar 4. Tampilan Form Hasil Prediksi

Keterangan:

Gambar 4 adalah hasil prediksi dari proses data yang telah dimasukkan ke sistem sebelumnya, sehingga memperoleh suatu nilai estimasi untuk besaran pendapatan dinas pendapatan daerah kabupaten deli serdang.

## **5. KESIMPULAN**

Berdasarkan hasil akhir pada penelitian skripsi ini penulis menarik sebuah Kesimpulan yang sekiranya dapat bermanfaat bagi pembaca, adapun kesimpulannya sebagai berikut:

1. Berdasarkan proses hasil perhitungan yang telah dilakukan, maka dapat diketahui nilai prediksi besaran pendapatan daerah dinas pendapatan daerah kab. deli serdang.
2. Dengan menggunakan algoritma regresi linear sederhana, dinilai dapat memprediksi besaran pendapatan daerah dinas pendapatan daerah kab. deli serdang untuk tahun berikutnya sehingga program-program yang telah direncanakan sebelumnya oleh dinas pendapatan dapat berjalan dengan lancar, dan juga dapat membuat program-program yang baru agar dapat meningkatkan pendapatan daerah untuk memajukan daerah tersebut.
3. Perancangan aplikasi dalam memprediksi besaran pendapatan daerah dinas pendapatan daerah kab. deli serdang dilakukan dengan menggunakan *visual basic net 2008*.

## **6. REFERENCES**

- [1] M. Marbun, H. T. Sihotang and M. A. Nababan, "Perancangan Sistem Peramalan Jumlah Wisata Asing," vol. 2, no. 1, p. 41, 2018.
- [2] D. P. S. P. M. F. U. Anwar, "Mendapatkan Algoritma Menghitung Koefisien-koefisien Persamaan Regresi Linear Dua Variabel untuk Dikoding ke dalam Bahasa Pemrograman," vol. 1, no. 1, p. 9, 2012.
- [3] K. & E. T. Luthfi, *Algoritma dan Data Mining*, Yogyakarta: Penerbit Andi, 2009.
- [4] D. N. a. G. w. Nurcahyo, *Algoritma Data Mining Dan Pengujian*, Jakarta: Deepublish, 2015.
- [5] H. T. S. M. A. N. Murni Marbun, "PERANCANGAN SISTEM PERAMALAN JUMLAH WISATAWAN ASING," vol. 2, no. 1, p. 42, 2018.
- [6] S. H. P. D. R. M. H. M. Mandala Harefa, *Optimalisasi Kebijakan Penerimaan Daerah*, Jakarta: YayasanPustaka Obor Indonesia, 2017.
- [7] A. T. M. N. W. S. L. Achmad Sani Alhusain, *Kebijakan dan Strategi Peningkatan Pendapatan Asli Daerah dalam Pembangunan Nasional*, Jakarta: Yayasan Pustaka Obor Indonesia, 2017.
- [8] M. Drs. Lamhot Sitorus, *Algoritma Dan pemograman*, Yogyakarta: Andi, 2015.
- [9] Krismiaji, *Sistem Informasi Akuntansi*, Yogyakarta: UPP AMP YKPN, 2010.
- [10] B. Soeherman and M. Pinontoan, *Designing Information System*, Yogyakarta: Elex Media Komputindo, 2008.
- [11] R. A. Sukanto and M. Shalahuddin, *Rekayasa Perangkat Lunak (Terstruktur dan Berorientasi Objek)*, Informatika: Bandung, 2016.
- [12] Hendrayudi, *VB 2008 untuk Berbagai Keperluan Programming*, Jakarta: PT. Elex Media Komputindo, 2009.
- [13] A. Saputra, "Manajemen Basis Data Mysql Pada Situs FTP Lapan Bandung," vol. 13, no. 4, pp. 155-162, 2012.