

PENERAPAN DATA MINING UNTUK PREDIKSI NILAI TUKAR PETANI TANAMAN PANGAN DI INDONESIA DENGAN METODE LINEAR REGRESSION DAN SUPPORT VECTOR MACHINE

Kennardi¹⁾ dan Akhmad Budi²⁾

¹⁾Alumni Program Studi Sistem Informasi

²⁾ Staff Pengajar Program Studi Teknik Informatika
Institut Bisnis dan Informatika Kwik Kian Gie
Jl. Yos Sudarso Kav.87 Sunter Jakarta Utara 14350

<http://www.kwikkiangie.ac.id>

akhmad.budi@kwikkiangie.ac.id

ABSTRACT

Development of technology and information systems use data mining as a method of data processing is growing, due to the large amount of data available and growing every day, especially agricultural data. One of the benefits that can be obtained by applying information technology, among others, is to conduct the analysis and prediction of large amounts of data. The exchange value of food crop farmers (NTPP) is a factor of economic welfare measurement in Indonesia, but quite difficult to predict because there has not been a proper measurement prediction method for predicting the data (NTPP) in the future. This research will be carried out to measure the accuracy of data prediction results using the Linear Regression and Support Vector Machine, with the aim of obtaining the results of the level of accuracy offered by the two methods.

Key Words: *Prediction, Farmers Exchange Rate, Food Crop Farmers Exchange Rate, Data Mining, Linear Regression, Support Vector Machine, Rapidminer.*

1. PENDAHULUAN

Perkembangan teknologi informasi di era globalisasi telah menjadi hal yang tidak dapat dipisahkan seiring dengan berkembangnya kehidupan manusia. Semakin bertambahnya jumlah dan kebutuhan pengolahan data di dunia mengakibatkan perkembangan teknologi pengolahan data juga semakin pesat karena permintaan jumlah data yang akan diolah semakin bertambah. Kejadian yang terjadi setiap waktu dalam kehidupan manusia dapat diterjemahkan sebagai sebuah fakta, dimana fakta tersebut kemudian diubah menjadi sebuah data yang kemudian diproses menjadi sebuah informasi yang memiliki bentuk atau pola yang mampu digunakan oleh sebuah sistem. Sehingga untuk pengolahan data dalam jumlah yang besar dibutuhkan sistem atau metode yang dapat menghasilkan informasi yang akurat secara efektif dan efisien.

Dalam perkembangannya di berbagai bidang, teknologi informasi dibuat agar dapat mempermudah dan mempercepat kinerja manusia dalam kehidupan sehari-hari dalam

menyelesaikan berbagai masalah yang dihadapi. Sehingga sistem pengolahan data dituntut untuk terus beradaptasi dan mampu menyelesaikan masalah secara cepat dan efektif. Salah satu manfaat yang bisa didapat dengan menerapkan teknologi informasi antara lain adalah melakukan analisis dan prediksi data dalam jumlah yang besar. Maka pada kesempatan ini penulis melakukan penelitian dengan mengukur keakuratan prediksi data dengan sampel nilai tukar petani tanaman pangan (NTPP) per provinsi di Indonesia.

Menurut *wikipedia*, Indonesia tercatat sebagai negara kepulauan terbesar di dunia dengan luas wilayah sebesar 1.904.569 km² dan memiliki jumlah penduduk terbesar ke-4 di dunia yaitu sebanyak 259.966.894 jiwa. Karena hal tersebut maka tidak dapat dipungkiri lagi bahwa data dan informasi yang dapat diperoleh dari negara Indonesia sangatlah besar jumlahnya jika dihitung dari berbagai faktor dan bidang yang saling berkembang di dalamnya. Salah satu bidang yang memiliki data dengan jumlah yang besar di Indonesia adalah bidang pertanian.

Sektor pertanian merupakan salah satu sektor yang sangat berkembang di Indonesia. Perkembangan ini dilandaskan karena Indonesia merupakan negara dengan iklim tropis yang sangat mendukung kegiatan bercocok tanam. Karena hal tersebut maka kebutuhan untuk memproses data pertanian di Indonesia dilakukan secara berkala untuk menemukan data indeks nilai tukar petani (NTP).

Data indeks nilai tukar pertanian di Indonesia dikumpulkan dan dihitung melalui Badan Pusat Statistik (BPS) yaitu lembaga pemerintah non kementerian yang bertanggung jawab langsung kepada presiden untuk mengumpulkan data tentang berbagai bidang di Indonesia untuk kebutuhan pemerintah dan publik. Data yang dikumpulkan dari BPS tersebut dapat diakses melalui situs resmi BPS yaitu <http://bps.go.id> atau melalui portal data publik indonesia yaitu <http://data.go.id>. Data indeks NTP yang dikumpulkan oleh BPS memiliki keakuratan yang tinggi karena dikumpulkan berdasarkan data tiap provinsi yang dilakukan dengan melakukan observasi langsung ke lapangan dan menggunakan tenaga ahli statistik dalam pengumpulan dan penghitungannya

Nilai tukar petani (NTP) merupakan salah satu indikator yang sangat penting karena mengukur seberapa besar tingkat kesejahteraan seorang petani akan hasil tanaman yang dikelolanya. Dengan tersedianya banyak jenis tanaman dalam kegiatan pertanian di Indonesia maka penelitian penulis kali ini berfokus pada indeks nilai tukar tanaman pangan (NTPP) per provinsi di Indonesia. Indonesia sebagai negara dengan jumlah penduduk yang besar memiliki kebutuhan pangan yang sangat besar pula sehingga kebutuhan dan kualitas tanaman pangan perlu diperhatikan agar pengelolaan dan distribusinya dapat terjangkau dan tertata dengan baik.

Pengelolaan sumber daya pertanian terutama tanaman pangan sangat bergantung dengan nilai tukar petani tanaman pangan (NTPP) yang menunjukkan indeks kesejahteraan petani, sehingga jika angka NTPP menunjukkan angka yang rendah maka peningkatan kualitas tanaman pangan di

daerah tersebut perlu ditingkatkan agar distribusi secara nasional dapat tercapai secara efektif dan tidak terjadi penyimpangan jumlah tanaman pangan. Penghitungan NTPP secara umum dilakukan dengan cara membandingkan IT (indeks diterima) dan IB (indeks dibayar).

Tentu saja dalam pengelolaan data pertanian dan penghitungan NTPP ada beberapa kesulitan yang muncul, yaitu antara lain adalah kesulitan memprediksi nilai NTPP menggunakan metode konvensional karena data IT dan IB diperoleh berdasarkan survey dan pendataan langsung yang memakan banyak waktu, sehingga penghitungan prediksi yang dilakukan harus menunggu sampai pendataan nasional selesai dilakukan. Serta hal tersebut menyebabkan tidak diketahuinya data *testing* dalam periode tertentu yang paling akurat untuk memprediksi hasil NTPP di masa depan.

Masalah lain yang timbul adalah karena belum adanya metode prediksi yang digunakan untuk memprediksi nilai NTPP selain menunggu hasil survey dan pencatatan metode konvensional yang biasa digunakan BPS.

Oleh karena itu dalam penelitian karya akhir ini penulis melakukan penelitian dengan cara memprediksi jumlah nilai NTPP dengan menggunakan metode *Linear Regression* dan *Support Vector Machine* untuk mengukur metode yang paling akurat untuk digunakan sebagai dasar metode prediksi NTPP di masa depan jika dibutuhkan. Hal ini juga dilandasi karena belum adanya penelitian pengukuran akurasi prediksi NTPP untuk membandingkan hasil prediksi metode *Linear Regression* dan *Support Vector Machine*.

2. LANDASAN TEORI

2.1 Sistem Informasi

Sistem informasi dapat merupakan kombinasi teratur apa pun dari orang-orang, *hardware*, *software*, jaringan komunikasi, dan sumber daya data yang mengumpulkan, mengubah, dan menyebarkan informasi, dalam sebuah organisasi [10]. Pendapat lain mengatakan bahwa sistem informasi adalah seperangkat unsur yang saling terkait atau komponen yang mengumpulkan (input),

memanipulasi (proses), menyimpan, dan menyebarkan (output) data dan informasi dan memberikan reaksi (mekanisme umpan balik) korektif untuk memenuhi tujuan [9].

2.2 Data

Data adalah sebuah fakta mentah [9] atau aliran fakta mentah yang mewakili peristiwa yang terjadi dalam organisasi atau lingkungan fisik sebelum diatur dan disusun menjadi bentuk yang dapat dipahami dan digunakan [8].

2.3 Database

Database merupakan koleksi terorganisir dari fakta-fakta dan informasi, biasanya terdiri dari dua atau lebih terkait file data [9].

2.4 Data Mining

Menurut [6] *data mining* adalah proses menemukan pola yang menarik dan pengetahuan dari sejumlah besar data atau *data mining* adalah istilah yang mengacu pada penggunaan algoritma dan komputer untuk menemukan novel dan pola menarik dalam data [11].

2.5 Jenis-jenis Data Mining

Menurut [6] ada beberapa jenis fungsionalitas *database* yang dapat yang dapat dilakukan dalam pengolahan data antara lain adalah :

1. *Class/Concept Description: Characterization and Discrimination* (Deskripsi kelas/konsep : karakterisasi dan diskriminasi) :
Entri data yang dapat dikaitkan dengan kelas atau konsep. Fungsi ini terdiri dari :
 - a. *Data Characterization* (Karakterisasi Data) : adalah ringkasan dari karakteristik umum atau fitur dari kelas target data.
 - b. *Data Discrimination* (Diskriminasi Data) : adalah perbandingan fitur umum dari objek data kelas sasaran terhadap fitur umum objek dari satu atau beberapa kelas yang kontras.

2. *Mining Frequent Patterns Associations, and Correlations* (Penggalian pola yang sering muncul : asosiasi dan korelasi) : Meneliti pola yang sering terjadi di data. Fungsi terdiri dari :
 - a. *Associations* (Asosiasi) : Pola dimana suatu variabel memiliki confidence (tingkat keyakinan) dengan variabel lain dan support (tingkat pendukung) dimana variabel memiliki pola yang sama.
 - b. *Correlations* (Korelasi) : Tingkat hubungan yang dimiliki oleh suatu variabel dengan variabel lain.
3. *Classifications and Regression for Predictive Analysis* (Klasifikasi dan regresi untuk analisis prediksi) :
 - a. *Classification* (Klasifikasi) : adalah proses menemukan model (atau fungsi) yang menggambarkan dan membedakan kelas data atau konsep.
 - b. *Regression* (Regresi) : adalah proses untuk mengestimasi hubungan antara variabel.
4. *Cluster analysis* (Analisis kluster) : pengelompokan analisis objek data tanpa konsultasi label kelas.

Sedangkan menurut [8] ada beberapa macam informasi yang dapat kita dapatkan dengan melakukan pengolahan data dalam *data mining* yaitu antara lain :

1. *Association/Asosiasi* : *Association* adalah kejadian terkait dengan peristiwa tunggal. Misalnya, sebuah studi dari pola pembelian supermarket mungkin mengungkapkan bahwa, ketika chip jagung dibeli, minuman cola dibeli sebanyak 65 persen dari waktu ke waktu tetapi ketika ada promosi, cola dibeli sebanyak 85 persen dari waktu ke waktu . Informasi ini membantu manajer membuat keputusan yang lebih baik karena mereka telah belajar profitabilitas promosi.
2. *Sequences/Pengurutan* : *Sequences* adalah peristiwa terkait dari waktu ke waktu. Kita mungkin menemukan, misalnya, bahwa jika seseorang membeli rumah, maka kulkas baru akan dibeli dalam waktu dua minggu

sebanyak 65 persen dari waktu ke waktu dan oven akan dibeli dalam waktu satu bulan dari pembelian rumah sebanyak 45 persen dari waktu ke waktu .

3. *Classification*/klasifikasi : *Classification* mengakui pola yang menggambarkan kelompok yang dimiliki oleh sebuah item dengan memeriksa item yang ada yang telah diklasifikasikan dengan menyimpulkan seperangkat aturan. Misalnya, bisnis seperti kartu kredit atau telepon perusahaan khawatir tentang hilangnya pelanggan tetap. Klasifikasi membantu menemukan karakteristik pelanggan yang cenderung untuk pergi dan dapat memberikan model untuk membantu manajer memprediksi siapa orang pelanggan sehingga manajer dapat merancang kampanye khusus untuk mempertahankan pelanggan tersebut.
4. *Clustering*/segmentasi : *Clustering* bekerja dalam cara yang mirip dengan klasifikasi ketika ada kelompok yang belum ditentukan. Sebuah alat data mining dapat menemukan kelompok yang berbeda dalam data, seperti menemukan kelompok afinitas untuk kartu bank atau partisi database ke dalam kelompok pelanggan berdasarkan demografi dan jenis investasi pribadi
5. *Forecasting/Prediction*/prediksi : *Forecasting* menggunakan prediksi dengan cara yang berbeda. Menggunakan serangkaian nilai-nilai yang ada untuk meramalkan nilai-nilai lainnya yang akan terjadi. Misalnya, peramalan mungkin menemukan pola dalam data untuk membantu manajer memperkirakan nilai masa depan dari variabel kontinu, seperti angka penjualan.

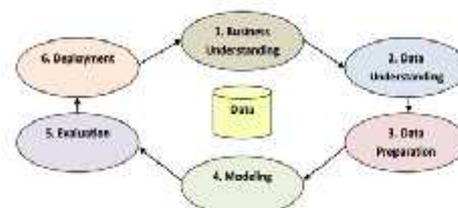
2.6 Tahap-tahap Data Mining

Tahapan dalam *data mining* terbagi dalam beberapa langkah yang disebut CRISP-DM (CRoss-Industry Standard Process for Data Mining) [1] yaitu antara lain adalah :

1. Business Understanding/Organizational

Understanding (Pemahaman bisnis/organisasi) : Tahap pemahaman sistem yang berjalan dan kebutuhan apa yang dibutuhkan dalam menyelesaikan masalah yang timbul didalamnya.

2. Data Understanding (Pemahaman data) : Tahap pemahaman dan pengumpulan data yang dibutuhkan untuk sebelum dilakukan persiapan untuk analisa. Pada tahap ini data yang dikumpulkan harus merupakan data yang tepat digunakan untuk proses penelitian dan mewakili masalah yang akan dipecahkan serta sesuai dengan kebutuhan dan kepentingan.
3. Data Preparation (Persiapan data) : Tahap persiapan dan seleksi data yang telah dikumpulkan dan diubah menjadi bentuk yang dapat diolah dalam model yang ditentukan selanjutnya.
4. Modeling (Pemodelan) : Proses analisa dan pemodelan data yang telah disiapkan dimana dalam ini dilakukan penerapan atau penghitungan berdasarkan algoritma atau metode yang ditentukan untuk mendapatkan hasil yang diinginkan sesuai dengan kebutuhan pengguna dan melakukan representasi pemecahan masalah.
5. Evaluation (Evaluasi) : Melakukan analisa dan evaluasi dari hasil model yang telah dibuat apakah sudah sesuai standar dan telah memecahkan masalah atau memenuhi kebutuhan dari pengguna.
6. Deployment (Penerapan) : Tahap penerapan hasil dari model yang telah dievaluasi dan dianalisa untuk kemudian dijadikan bentuk yang dapat diolah kembali.



Gambar 1. Tahapan data mining (CRISP-DM)

Dalam penerapannya data mining adalah sebuah siklus yang terbagi dalam 6 tahap yang disebut CRISP-DM (CRoss-

Industry Standard Process for Data Mining [5] yaitu :

1. *Business Understanding* (Pemahaman bisnis) :
 - a. *Determining Business Objectives* (Menentukan tujuan bisnis).
 - b. *Assess Situation* (Menilai situasi).
 - c. *Determining Data Mining Goals* (Menentukan tujuan penambangan data).
 - d. *Produce Project Plan* (Membuat rencana proyek).
2. *Data Understanding* (Pemahaman data) :
 - a. *Collect Initial Data* (Mengumpulkan data utama).
 - b. *Describe Data* (Mendeskripsikan data).
 - c. *Explore Data* (Menjelajah data).
 - d. *Verfify Data Quality* (Verifikasi kualitas data).
3. *Data Preparation* (Persiapan data) :
 - a. *Select Data* (Memilih data).
 - b. *Clean Data* (Pembersihan data).
 - c. *Construct Data* (Membangun/membentuk data).
 - d. *Intergrate Data* (Menggabungkan data).
 - e. *Format Data* (Memformat data).
4. *Modeling* (Pemodelan) :
 - a. *Select Model Techniques* (Memilih teknik pemodelan data).
 - b. *Generate Test Design* (Menghasilkan uji desain).
 - c. *Build Model* (Membangun model).
 - d. *Assess Model* (Evaluasi/penilaian model).
5. *Evaluation* (Evaluasi) :
 - a. *Evaluate Results* (Evaluasi hasil).
 - b. *Review Process* (Tinjauan ulang proses).
 - c. *Determine Next Steps* (Penentuan langkah selanjutnya).
6. *Deployment* (Penerapan)
 - a. *Deployment Plan* (Penerapan rencana).

- b. *Plan Monitoring and Maintenance* (Pemantauan dan pemeliharaan rencana).
- c. *Produce Final Report* (Menghasilkan laporan akhir).
- d. *Review Project* (Tinjauan ulang proyek).

2.7 Linear Regression

Menurut [3], *Linear Regression* adalah menemukan persamaan garis yang berhubungan antara satu variabel dengan yang lain. *Data mining* menggunakan persamaan ini bertujuan untuk memprediksi nilai satu variabel berdasarkan nilai lain.

2.8 Support Vector Machine

Support Vector Machine (SVM) adalah metode klasifikasi berdasarkan diskriminan marjin linear maksimum, yaitu, tujuannya adalah untuk menemukan *hyperplane* optimal yang memaksimalkan celah atau marjin antara kelas [7].

3. METODOLOGI PENELITIAN

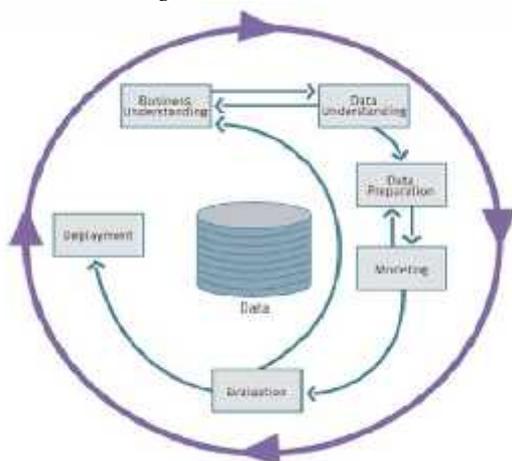
3.1 Teknik Pengumpulan Data

Dalam penelitian ini data yang digunakan adalah data sekunder dari nilai tukar petani tanaman pangan yang diambil dari <http://data.go.id/> berdasarkan survey dan pendataan yang dilakukan oleh Badan Pusat Statistik (BPS).

Dalam pengumpulan data penulis menggunakan metode kuantitatif yaitu penelitian ilmiah yang sistematis terhadap bagian-bagian dan fenomena serta hubungan-hubungannya. Tujuan penelitian kuantitatif adalah mengembangkan dan menggunakan model-model matematis, teori-teori dan/atau hipotesis yang berkaitan dengan fenomena alam. Proses pengukuran adalah bagian yang sentral dalam penelitian kuantitatif karena hal ini memberikan hubungan yang fundamental antara pengamatan empiris dan ekspresi matematis dari hubungan-hubungan kuantitatif.

3.2 Teknik Analisis Data

Teknik analisis data yang digunakan penulis dalam penelitian ini adalah teknik yang didasarkan pada metode CRISP-DM (*Cross-Industry Standard Process for Data Mining*).



Gambar 2. *Data mining cycle (CRISP-DM)*

3.3 Teknik Pengukuran

Pengukuran Akurasi Metode Prediksi :

Pengukuran kelayakan metode prediksi penulis menggunakan rumus Root Mean Squared Error untuk menentukan tingkat kesalahan prediksi data berdasarkan pengaruh dari tiap variabel :

$$RMSE = \sqrt{\frac{\sum_{t=1}^n (\hat{y}_t - y)^2}{n}}$$

\hat{y}_t = Data Prediksi

y = Data Aktual

n = Jumlah Data

Pengukuran Penyimpangan Prediksi :

Untuk mengukur penyimpangan data prediksi penulis melakukan pencarian rata-rata kesalahan (*Mean Error*), rata-rata kesalahan absolut (*Mean Absolute Error*), rata-rata persentase kesalahan (*Mean Percentage Error*), dan rata-rata persentase kesalahan absolut (*Mean Absolute Percentage Error*).

Mean Error (ME) :

$$ME = \frac{\sum E}{n}$$

E = Data Aktual – Data Prediksi
 n = Jumlah Data

Mean Absolute Error (MAE) :

$$MAE = \left| \frac{\sum E}{n} \right|$$

E = Data Aktual – Data Prediksi
 n = Jumlah Data

Mean Percentage Error (MPE) :

$$MPE = \frac{\sum \left(\frac{E}{A}\right) \times 100\%}{n}$$

E = Data Aktual – Data Prediksi
 n = Jumlah Data
 A = Data Aktual

Mean Absolute Percentage Error (MAPE) :

$$MAPE = \left| \frac{\sum \left(\frac{E}{A}\right) \times 100\%}{n} \right|$$

E = Data Aktual – Data Prediksi
 n = Jumlah Data
 A = Data Aktual

3.4 Teknik Perancangan GUI

Perancangan dan pengembangan GUI akan dilakukan dengan terlebih dahulu melakukan analisa data dari tahun 2008-2013 yang diperoleh sebagai data *training* dan data tahun 2014 sebagai data *testing* untuk dilakukan proses prediksi dalam *RapidMiner*. Sehingga diperoleh data prediksi dan aktual hasil menggunakan metode *Linear Regression* dan *Support Vector Machine*.

Kemudian penulis akan mengimplementasikan hasil prediksi dan analisa tersebut kedalam *Microsoft Access 2013* untuk kemudian melakukan perancangan GUI sederhana yang didalamnya terdapat proses penghitungan penyimpangan ke-

akuratan prediksi dan menghasilkan laporan komparasi data prediksi dan aktual secara sistematis dan siap diberikan pada pembaca untuk diteliti dan dipelajari lebih lanjut sesuai dengan metode yang dilakukan oleh penulis.

4. HASIL DAN PEMBAHASAN

4.1 Rancangan Sistem

Penulis melakukan perancangan sistem informasi berupa GUI (Graphic User Interface). Rancangan ini diperuntukkan untuk para pembaca dan pengguna aplikasi untuk memudahkan proses pengujian keakuratan prediksi NTPP (Nilai Tukar Tanaman Pangan) menggunakan *Data Mining* dengan metode *Linear Regression* dan *Support Vector Machine* dan menghitung seberapa besar *error margin* dari hasil pengujian tersebut. Perancangan GUI dilakukan dengan menggunakan aplikasi *Microsoft Access 2013*.

4.2 Rancangan Basis Data

Sebelum implementasi GUI dilakukan, akan dilakukan proses prediksi data untuk memperoleh data prediksi dengan menggunakan algoritma *Linear Regression* dan *Support Vector Machine*. Penulis meng-

gunakan aplikasi *Rapidminer 5.3* sebagai aplikasi pengolahan dan prediksi data yang dibutuhkan. Proses tersebut terdiri dari :

1. Pengujian Data Prediksi :
 - a. *Import Dataset* (Memasukkan Dataset)
 - b. *Process Building* (Pembangunan Alur Proses):
2. Pengujian Data Prediksi :
3. Pengujian *Error Margin* (Penyimpangan Prediksi)

4.3 Rancangan Basis Data

Setelah data dan pengukuran hasil prediksi didapat tahap selanjutnya adalah menerapkan data dan hasil analisa tersebut ke dalam GUI (Graphical User Interface) untuk memudahkan pembaca melihat hasil penelitian yang telah dilakukan. Perancangan GUI yang dilakukan menggunakan aplikasi *Microsoft Access 2013* untuk menghasilkan tampilan GUI yang simpel dan mudah dilihat oleh pembaca. Dalam penerapan sistem GUI penulis melakukan beberapa langkah yaitu :

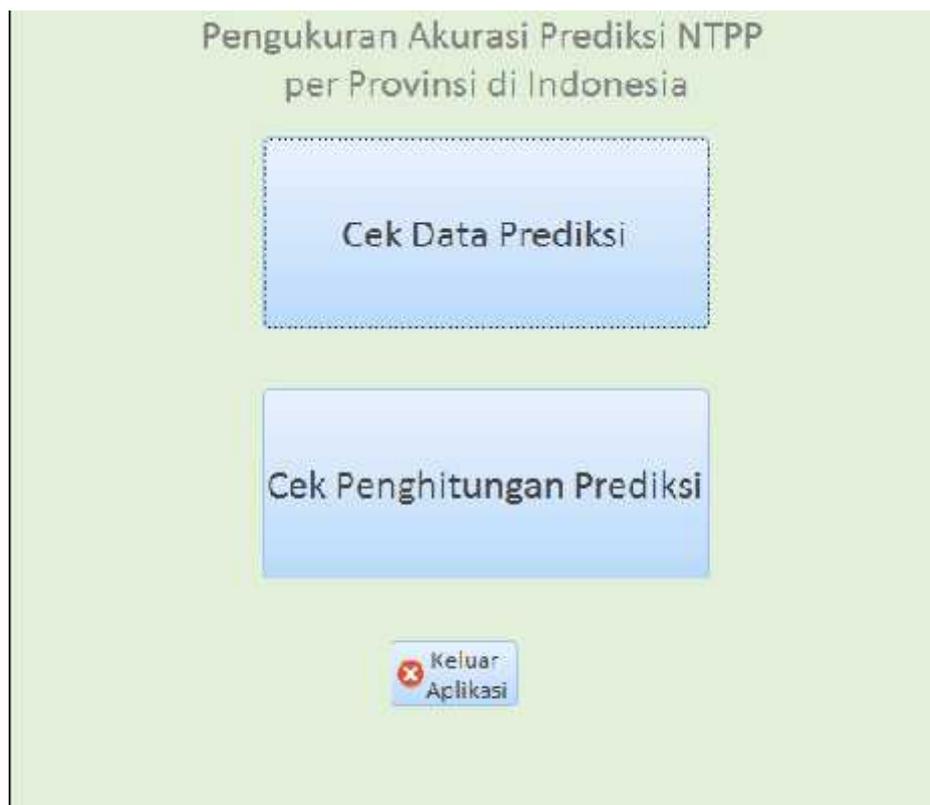
1. *Input* Data Prediksi
2. Proses Implementasi GUI
3. Spesifikasi Sistem

Tabel 1
Perbandingan *root_mean_squared_error*

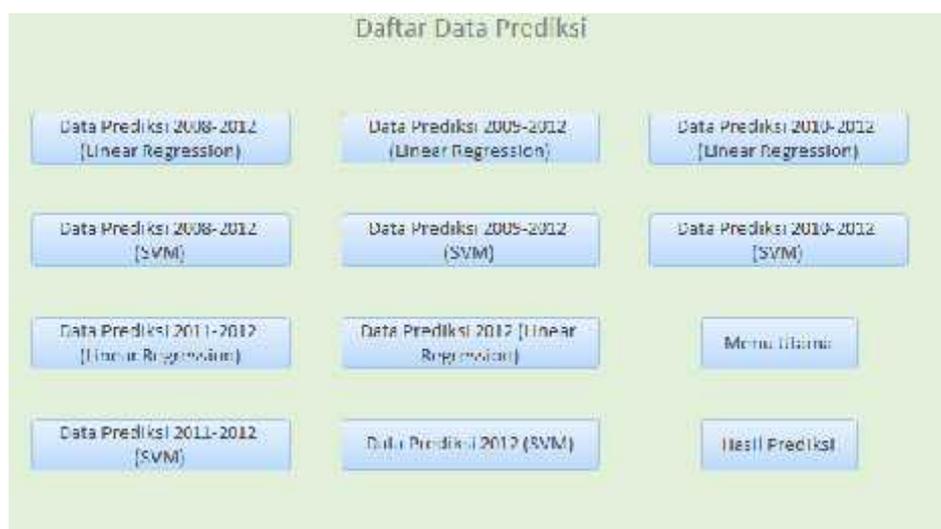
<i>Dataset</i>	Metode Prediksi	
	<i>Linear Regression</i>	SVM
2008-2012 (5 Tahun)	root_mean_squared_error 0.667 +/- 0.057 (mikro: 0.670 +/- 0.000)	root_mean_squared_error 0.676 +/- 0.063 (mikro: 0.679 +/- 0.000)
2009-2012 (4 Tahun)	root_mean_squared_error 0.603 +/- 0.051 (mikro: 0.605 +/- 0.000)	root_mean_squared_error 0.616 +/- 0.052 (mikro: 0.618 +/- 0.000)
2010-2012 (3 Tahun)	root_mean_squared_error 0.562 +/- 0.046 (mikro: 0.564 +/- 0.000)	root_mean_squared_error 0.590 +/- 0.064 (mikro: 0.594 +/- 0.000)
2011-2012 (2 Tahun)	root_mean_squared_error 0.507 +/- 0.071 (mikro: 0.512 +/- 0.000)	root_mean_squared_error 0.524 +/- 0.106 (mikro: 0.535 +/- 0.000)
2012 (1 Tahun)	root_mean_squared_error 0.485 +/- 0.062 (mikro: 0.489 +/- 0.000)	root_mean_squared_error 0.508 +/- 0.078 (mikro: 0.514 +/- 0.000)

Tabel 2
Perbandingan Tingkat Akurasi

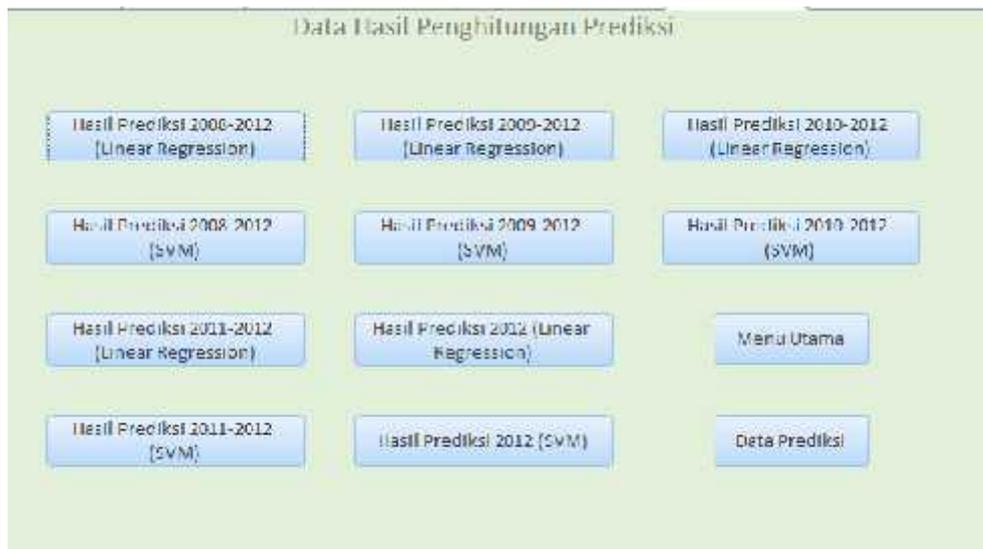
<i>Dataset</i>	Metode Prediksi	
	<i>Linear Regression</i>	SVM
2008-2012 (5 Tahun)	ME: -0.012900 MAE : 1.318584 MPE : 0.185549 % MAPE : 1.356569 %	ME : -0.134426 MAE : 1.211442 MPE : 0.040930 % MAPE : 1.233014 %
2009-2012 (4 Tahun)	ME : 0.008110 MAE : 1.198826 MPE : 0.187115 % MAPE : 1.235498 %	ME : -0.113280 MAE : 1.087869 MPE : 0.041696 % MAPE : 1.107403 %
2010-2012 (3 Tahun)	ME : 0.085838 MAE : 1.074489 MPE : 0.245976 % MAPE : 1.115297 %	ME : 0.013348 MAE : 0.95016 MPE : 0.149942 % MAPE : 0.979364 %
2011-2012 (2 Tahun)	ME : 0.033387 MAE : 0.911892 MPE : 0.163237 % MAPE : 0.943717 %	ME : 0.047569 MAE : 0.794807 MPE : 0.154469 % MAPE : 0.828507 %
2012 (1 Tahun)	ME : 0.037694 MAE : 0.756582 MPE : 0.136503 % MAPE : 0.789332 %	ME : 0.077018 MAE : 0.677481 MPE : 0.156788 % MAPE : 0.719025 %



Gambar 3. Form Menu Utama



Gambar 4. Form Sub Menu

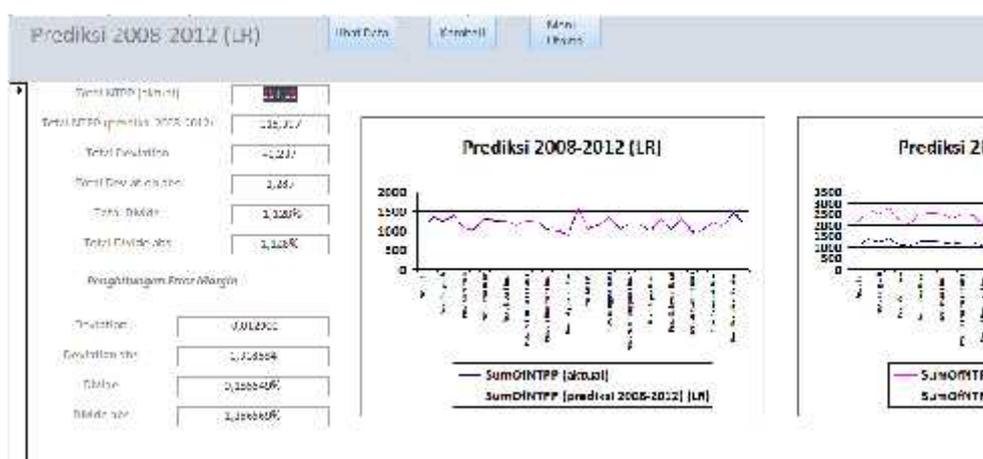


Gambar 5. Form Sub Menu 2

Prediksi-2008-2012 (LR)

Tahun	Index	nama Presiden	Kawab Presiden	IT	IS	NTTP (actual)	NTTP (predik)
2012	1	11. Provs. Nanggroe Aceh Darussalam		190,64	31,32	12,1	
2014	1	14. Provs. Sumatera Barat		140,99	21,36	16,44	
2012	1	15. Provs. Sumatera Barat		150,30	24,88	24,79	
2012	1	16. Provs. Riau		140,04	18,30	-12,13	
2012	1	19. Provs. Jambi		122,80	17,15	22,21	
2014	1	20. Provs. Sumatera Selatan		140,30	18,3	-12,35	
2011	1	17. Provs. Bengkulu		161,46	20,12	10,4	
2011	1	18. Provs. Lampung		101,76	17,01	16,55	
2011	1	19. Provs. Kepulauan Bangka Belitung		110,21	17,01	11,22	
2012	1	21. Provs. Kepulauan Riau		80,77	11,1	11,7	
2012	1	22. Provs. Irian Barat		101,27	15,3	10,1	

Gambar 6. Report Data Prediksi



Gambar 7 Form data penghitungan prediksi

4.4 Implikasi Penelitian

Penelitian ini memberikan implikasi bagi institusi yaitu dalam penelitian ini adalah BPS (Badan Pusat Statistik), antara lain :

1. Menyediakan hasil penelitian analisa dan prediksi NTPP selain menggunakan metode konvensional.
2. Menerapkan metode prediksi NTPP menggunakan metode *Linear Regression* dan *Support Vector Machine*.
3. Menyediakan hasil laporan perbandingan antara data prediksi dan data aktual untuk dijadikan pertimbangan dalam pengolahan data NTPP di masa depan.

5. SIMPULAN

Berdasarkan penelitian mengukur keakuratan prediksi NTPP menggunakan metode *Linear Regression* dan *Support Vector Machine* penulis menyimpulkan hal-hal berikut :

1. Hasil penelitian memberikan kemudahan dalam memprediksi data karena proses prediksi dilakukan dengan menggunakan metode *Linear Regression* dan *Support Vector Machine* berdasarkan data IT, IB, dan NTPP tahun lalu, sehingga waktu yang dibutuhkan untuk memproses data prediksi relatif lebih singkat.
2. Diperolehnya data *testing* yang paling akurat untuk memprediksi nilai NTPP berdasarkan penelitian yang dilakukan selama 5 periode yaitu data prediksi dengan data tahun 2012 sebagai *predictor* untuk metode *Linear Regression* dan *Support Vector Machine*.
3. Penggunaan metode *Linear Regression* dan *Support Vector Machine* untuk memprediksi nilai NTPP yang terbukti akurat dalam menentukan nilai NTPP, Sehingga kedua metode tersebut layak dipertimbangkan sebagai metode prediksi selain menggunakan metode konvensional.

4. Dalam penelitian yang dilakukan hasil dari pengukuran prediksi akurasi, diperoleh fakta bahwa metode *Support Vector Machine* memiliki nilai *error margin* yang lebih kecil dibanding *Linear Regression* sehingga tingkat akurasi metode tersebut lebih cocok untuk digunakan.

6. REKOMENDASI

Dalam pembuatan penelitian karya akhir ini penulis sekiranya memberi saran lanjutan untuk para pembaca untuk dipertimbangkan antara lain :

1. Hasil data prediksi yang diperoleh oleh penulis dari penelitian ini dilakukan berdasarkan prediksi dari data dengan *record* yang besar untuk memprediksi *record* yang kecil, maka disarankan untuk menggunakan teknik sebaliknya untuk mendapatkan hasil referensi yang lebih tepat sebagai pembanding penelitian selanjutnya.
2. Penghitungan prediksi menggunakan *Support Vector Machine* masih sangat sulit jika dilakukan secara manual, karena bentuk algoritma yang rumit dibandingkan dengan *Linear Regression*. Jika ada penelitian selanjutnya yang menggunakan algoritma yang sama disarankan untuk mencantumkan penghitungan secara manual sebagai referensi hasil prediksi yang lebih akurat.
3. Rancangan GUI yang dibuat oleh penulis sengaja dibuat secara simpel agar mudah dimengerti dan tidak memakan waktu karena implementasi algoritma dilakukan dengan segera, maka bila dirasa kurang memuaskan dapat dikembangkan atau ditambahkan sesuai keinginan pengembang selanjutnya agar tampilan GUI dapat lebih diperbaiki.
4. Sebagaimana sebuah penelitian, metode dan hasil yang diperoleh dari penelitian ini tidaklah menjadi sebuah hal yang "baku", karena itu hasil penelitian ini dapat saja diperbaharui di masa depan oleh peneliti selanjutnya

agar dapat menemukan metode dan hasil yang lebih efektif.

7. DAFTAR PUSTAKA

- [1] A.North, Matthew (2012), *Data Mining for the Masses*, Edisi ke-1, Georgia : Global Text Project.
- [2] Apriyanti Ira (2003), Skripsi : Analisis Nilai Tukar Petani Di Sumatra Utara, Universitas Sumatra Utara, Medan.
- [3] Brown, Meta S. (2014), *Data Mining For Dummies*, New Jersey : John Wiley & Sons, Inc.
- [4] Carneiro da Rocha, Bruno (2010), *Identifying Bank Frauds Using CRISP-DM and Decision Trees*, International journal of computer science & information Technology (IJCSIT) Vol.2, No.5.
- [5] Erskine, Joseph R. (2010), Tesis : *Developing Cyberspace Data Understanding : Using CRISP-DM for Host-based IDS Feature Mining*, USAF, Ohio.
- [6] Han, Jiawei et al (2012), *Data Mining Concepts and Techniques*, Edisi ke-3, Waltham, Elsevier Inc.
- [7] J. Zaki, Mohammed dan Wagner Meira JR (2014), *Data Mining and Analysis Fundamentals Concepts and Algorithms*, New York : Cambridge University Press.
- [8] Laudon, Kenneth C. dan Jane P. Laudon, (2012), *Management Information Systems:Managing the Digital Firm*, Edisi ke-12, New Jersey : Pearson Prentice Hall.
- [9] M.Stair, Ralph dan George W. Reynolds (2012), *Fundamentals of Information Systems*, Edisi ke-6, Boston : Cengage Learning.
- [10] O'Brien, James A. (2010), *Management Information System*, Edisi ke-15, New York: Mc Graw Hill Irwin.
- [11] Stanton, Jeffrey dan Robert W. De Graaf , *Version 3 : An Introduction to Data Science*, New York : Creative Commons.
- [12] Suhendar Asep (2013), Skripsi : *Analisis Faktor-faktor yang Memengaruhi Nilai Tukar Petani Tanaman Pangan di Kawasan Barat Indonesia (Periode Tahun 2008-2010)*, Bogor Agricultural University, Bogor.