

Prediksi Angka Kemiskinan Di Kabupaten Ogan Komering Ulu Menggunakan Metode Data Mining Berbasis *SVM*, *Neural Network* & *SGD*

Daniel Kurniawan^{1*}, Silvia Apriani², Pujiyanto³

¹Informatika/Teknik dan Komputer
Universitas Baturaja

^{1*}danielkurniawan062@gmail.com

²Informatika/Teknik dan Komputer
Universitas Baturaja

²silviapriani28@email.com

³Informatika/Teknik dan Komputer
Universitas Baturaja

³Pujiyanto.mail@email.com

Abstrak— Kemiskinan adalah masalah kompleks yang dipengaruhi oleh berbagai faktor ekonomi, sosial, dan demografi yang saling berkaitan. Sudah menjadi tugas kebijakan ekonomi yang signifikan, oleh karena itu, dalam hal tersebut, untuk mengidentifikasi segala faktor yang relevan dan signifikan serta menguraikan efeknya terhadap tingkat kemiskinan, kita menggunakan data sekunder yang diperoleh dari Badan Pusat Statistik dan sumber-sumber kemiskinan lainnya yang memengaruhi periode penelitian yang spesifik. Penelitian ini bertujuan untuk mengeksplorasi penerapan Algoritma SVM (Support Vector Machine), Neural Network, Stochastic Gradient Descent. dengan menggunakan aplikasi Orange untuk memprediksi angka kemiskinan yang ada di Kabupaten OKU. Metodologi penelitian melibatkan tahapan pengumpulan data, pemrosesan, dan implementasi menggunakan metode SVM(Support Vector Machine), Neural Network, Stochastic Gradient Descent. melalui platform Orange. Hasil penelitian ini dapat memberikan wawasan mendalam mengenai kemampuan SVM(Support Vector Machine), Neural Network, Stochastic Gradient Descent. dalam memprediksi angka kemiskinan untuk hasil dari perhitungan yang sudah di dapatkan adalah Jumlah prediksi tahun 2025 = 72.42 sedangkan Jumlah prediksi tahun 2026 = 70.15 di Kabupaten OKU.

Kata kunci— Prediksi, Kemiskinan, Orange, Data Mining, SVM, Neural Network, Stochastic Gradient Descent..

Abstract— Poverty is a complex problem influenced by various interrelated economic, social, and demographic factors. It has become a significant economic policy task, therefore, in this case, to identify all relevant and significant factors and describe their effects on poverty levels, we use secondary data obtained from the Central Bureau of Statistics and other poverty sources that affect the specific research period. This study aims to explore the application of the SVM (Support Vector Machine) Algorithm, Neural Network, Stochastic Gradient Descent. using the Orange application to predict poverty rates in OKU Regency. The research methodology involves the stages of data collection, processing, and implementation using the SVM (Support Vector Machine) method, Neural Network, Stochastic Gradient Descent. through the Orange platform. The results of this study can provide in-depth insights into the capabilities of SVM (Support Vector Machine), Neural Network, Stochastic Gradient Descent. In predicting poverty figures, the results of the calculations obtained are the predicted number for 2025 = 72.42, while the predicted number for 2026 = 70.15 in OKU Regency.

Keywords— Prediction, Poverty, Orange, Data Mining, SVM, Neural Network, Stochastic Gradient Descent.

I. PENDAHULUAN

Salah satu isu utama yang dihadapi banyak negara berkembang, termasuk Indonesia, adalah kemiskinan. Di dibandingkan dengan negara maju, angka kemiskinan yang tinggi bukanlah indikasi utama kondisi ekonomi tidak merata namun adalah refleksi ketidakadilan dan kesenjangan sosial. Selain itu, hal ini juga berarti akses terbatas terhadap kebutuhan dasar seperti pendidikan, kesehatan, dan pekerjaan. Kemiskinan adalah problematika yang mengungkapkan bahwa seseorang tidak dapat

mencukupi kebutuhan dasar hidupnya sehari-sehari. Yang mana berarti ketidakberdayaan dalam mencapai hal dasar dalam hak hak yang seharusnya mereka terima [11].

Diperlukan kajian lebih mendalam, bahwa meskipun telah banyak program dan kebijakan digulirkan (strategi struktural) dengan dukungan anggaran yang tidak sedikit, penanggulangan kemiskinan masih belum berjalan dengan hasil yang optimal. Wajah kemiskinan yang dipengaruhi oleh banyak faktor (multidimensional), sangat dimungkinkan karena sebab-sebab kultural (strategi budaya)

masih menjadi salah satu bagian dari hambatan penanggulangan kemiskinan [5].

Pemerintah sendiri selalu mencanangkan upaya penanggulangan kemiskinan dari tahun ketahun, namun tingkat kemiskinan Indonesia tidak juga mengalami penurunan yang signifikan, walaupun data di BPS menunjukkan kecenderungan penurunannya, namun secara kualitatif belum menampakkan dampak perubahan yang nyata malahan kondisinya semakin memprihatinkan tiap tahunnya [6].

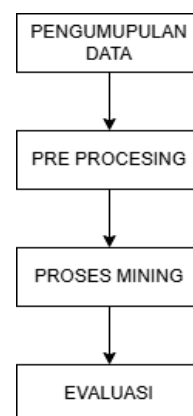
Dengan demikian, dalam upaya untuk lebih memahami dan meramal angka kemiskinan, ini penting untuk mengidentifikasi faktor-faktor ini. Mereka dapat dipasarkan sebagai faktor ekonomi, termasuk tingkat profit kerja dan inflasi faktor-faktor sosial ekspansi layanan, kapasitas pelayanan dan hubungan terfluktuasi faktor demografi distribusi penduduk, urbanisasi. Pemahaman lebih awal hubungan antara faktor-faktor ini dan kemiskinan akan membantu dalam perumusan kebijakan yang lebih tepat untuk menanganinya.

Metode SVM adalah metode orange yang digunakan untuk menganalisis hubungan antara satu variabel terikat dan beberapa variabel bebas. Dalam penelitian ini diterapkan untuk memodelkan dan memprediksi jumlah penduduk miskin di Indonesia berdasarkan beberapa faktor penentu. Dengan memanfaatkan data sekunder dari Badan Pusat Statistik dan sumber lainnya, penelitian ini bertujuan untuk mengidentifikasi faktor-faktor signifikan yang mempengaruhi tingkat kemiskinan dan menilai sejauh mana faktor-faktor tersebut dapat menjelaskan variasi dalam variabel terikat. Melalui analisis ini, diharapkan dapat ditemukan gambaran yang lebih jelas tentang faktor-faktor utama yang berkontribusi terhadap kemiskinan di Indonesia.

II. METODOLOGI PENELITIAN

Dalam penelitian ini pengolahan data menggunakan analisis prediktif dilakukan untuk melihat kondisi diwaktu yang akan datang dan berbagai kemungkinan yang dapat terjadi dengan teknik klasifikasi menggunakan SVM (Support Vector Machine) data mining.

Klasifikasi data mining mengelompokkan beberapa variabel dan diberikan label. Menggunakan bantuan metode Knowledge Discovery in Database (KDD) beberapa tahapan atau langkah untuk data mining seperti di bawah ini.



Gambar 1. *Discovery in Database (KDD)*

A. Pengumpulan Data

Variabel tersebut ada yang tergolong variable dependen yaitu variabel yang dijadikan sebagai variabel target adalah jumlah angka kemiskinan dan variabel independen yaitu variabel yang dijadikan sebagai faktor penyebab jumlah angka kemiskinan yang tidak beratur. Adapun variabel independennya adalah Tahun dalam penentuan jumlah kemiskinan tersebut yaitu dari tahun 2004 sampai dengan 2024. Sumber data di dapatkan Website Badan Pusat Statistik atau biasa di singkat BPS.

B. Pre Proccesing

Pada proses ini sampel data dibagi menjadi 2, yaitu data training 80% (tahun 2004–2020) dan data testing 20% (tahun 2021–2024, prediksi tahun 2025). Dalam penelitian ini data training dihasilkan data training 80% yaitu 17 dari 21 data dan data testing 20% yaitu 4 dari 21 data. Pada tahap ini dataset tidak mengalami missing value oleh karena itu dapat dilakukan tahap transformation. Pada penelitian ini tahap transformation yang dilakukan yaitu mengubah bentuk data, data yang masih berupa angka yang ditransformasikan kedalam bentuk kategorikal seperti pada variable jumlah penduduk.

C. Proses mining

Pada tahap ini yaitu melakukan penerapan teknik algoritma SVM. Proses ini adalah tahapan yang menentukan model pohon keputusan yang sesuai dengan data training[4]. Hasil penelitian dilakukan dengan menggunakan tools orange dengan jumlah data sebanyak 21 data jumlah penduduk yang miskin. Penggunaan data mining ini bertujuan untuk menentukan prediksi angka kemiskinan dengan menggunakan metode SVM (Support Vector Machine)[4].

D. Evaluasi

Tahap evaluasi dalam penelitian ini bertujuan untuk mengukur tingkat keakuratan model prediksi yang dibangun menggunakan algoritma Support Vector Machine (SVM). Metode ini mempartisi data ke dalam dua sub set data yang berukuran sama[14]. Model ini digunakan untuk memprediksi jumlah penduduk miskin (dalam ribuan jiwa) di Kabupaten OKU berdasarkan data historis dari tahun 2021 hingga 2024. Data tahun 2025 digunakan sebagai data uji.

Evaluasi dilakukan dengan menggunakan metrik Mean Absolute Percentage Error (MAPE) karena sifat target yang bersifat numerik dan kontinu. MAPE merupakan metrik yang umum digunakan dalam evaluasi model regresi untuk mengukur seberapa besar kesalahan prediksi dalam satuan persen relatif terhadap nilai aktual.

$$MAPE : \frac{1}{N} \sum_{t=1}^N \frac{|X_t - \hat{x}_t|}{\hat{x}_t} \times 100\%$$

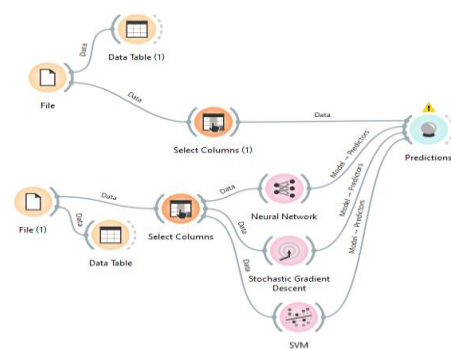
Dimana:

X_t : Hasil nilai observasi (Data rel)

\hat{x}_t : Hasil nilai peramalan ke-t

n : Banyak data ke-n

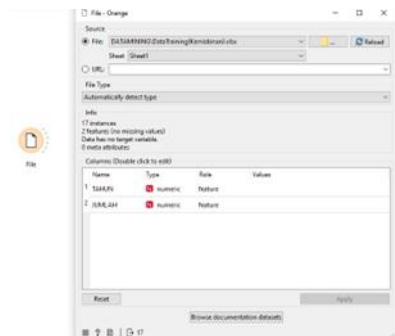
III. HASIL DAN PEMBAHASAN



Gambar 2. Model penelitian

Hasil penelitian dituliskan berdasarkan tahapan penelitian dari langkah awal penelitian hingga akhir seperti pada gambar 2.

A. Data File



Gambar 3. Widget File

Pada Gambar 3 melakukan input terhadap data hipertensi ke dalam widget file yang dimana datanya terdiri dari 20 instance dan 2 feature.

B. Data File

- Data Tabel Training

	TAHUN	JUMLAH
1	2004	201.40
2	2005	45.20
3	2006	46.10
4	2007	40.60
5	2008	38.60
6	2009	35.10
7	2010	39.90
8	2011	38.30
9	2012	37.60
10	2013	42.00
11	2014	41.40
12	2015	46.04
13	2016	46.97
14	2017	46.34
15	2018	45.71
16	2019	46.84
17	2020	47.30

Gambar 4. Data Tabel Training

Data table training tempat untuk melihat data yang sudah di input kedalam data file, Gambar 4 adalah bentuk data hipertensi yang digunakan dalam penelitian.

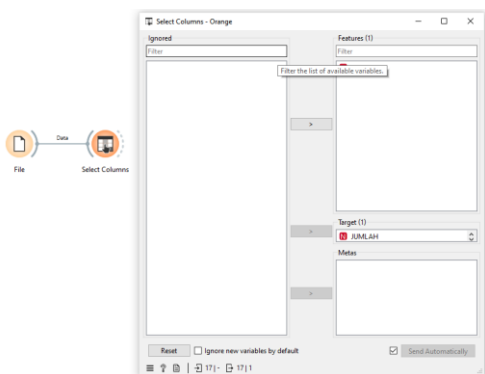
- Data Tabel Testing

	TAHUN	JUMLAH
1	2021	47.50
2	2022	44.20
3	2023	44.11
4	2024	41.54
5	2025	?
6	2026	?

Gambar 5. Data Tabel Testing

Data table testing tempat untuk melihat data yang sudah di input kedalam data file, Gambar 5 adalah bentuk data hipertensi yang digunakan dalam penelitian.

C. Data Column/Target

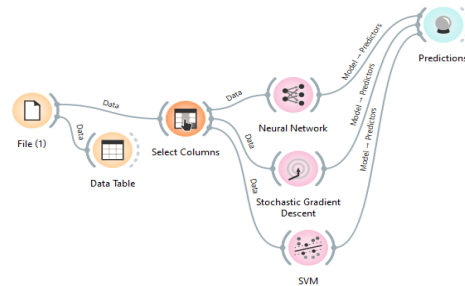


Gambar 6. Data Tabel Column

Columns akan dilakukan pengambilan satu features yang akan dijadikan sebagai target variable,

selanjutnya akan dilakukannya pengujian terhadap data tersebut menggunakan metode algoritma Support Vector Machine (SVM).

D. Algoritma Support Vector Machine (SVM), Neural Network, Stochastic Gradient Descent(SGD)



Gambar 7. Algoritma SVM Neural Network, Stochastic Gradient Descent(SGD).

Pada gambar 7 dilakukan penghubungan data dari select columns ke Support Vector Machine (SVM), Neural Network, Stochastic Gradient Descent(SGD). Setelah terhubung dengan Support Vector Machine (SVM) selanjutnya akan dihubungkan ke dalam proses predictions. Dalam proses predictions membutuhkan dua data yang sudah dilakukan perhitungan algoritma dan data yang belum dilakukan perhitungan algoritma oleh karena itu peneliti menghubungkan data dari select columns ke proses predictions untuk melihat hasilnya.

E. Predictions

	SVM	error	Stochastic Gradient Descent	error	Neural Network	error	JUMLAH	TAHUN
1	46.25	-1.25	28.14	-19.36	7.88	-39.62	47.50	2021
2	45.86	1.66	25.50	-18.70	8.39	-35.81	44.20	2022
3	45.46	1.35	22.87	-21.24	8.91	-35.20	44.11	2023
4	45.09	3.55	20.24	-21.30	9.46	-32.08	41.54	2024
5	44.79		17.60		10.03		?	2025
6	44.57		14.97		10.61		?	2026

Gambar 8. Hasil Data Predictions

Data hasil ini merupakan Hasil dari prediksi Algoritma Support Vector Descent (SVM), Neural Network dan Stochastic Gradient Descent.

F. Hasil Perhitungan MAPE

Model	MAPE
Stochastic Gradient Descent	0.456
SVM	0.045
Neural Network	0.804

Gambar 9. Data Hasil Perhitungan MAPE

Untuk menentukan hasil dari perhitungan MAPE tersebut harus mengetahui perumusannya terlebih dahulu yaitu sebagai berikut:

$$\text{MAPE} = \frac{1}{N} \sum_{i=1}^N \left| \frac{A_i - F_i}{A_i} \right| \times 100\%$$

Keterangan:

A_i : Hasil nilai observasi (Data rel)

F_i : Hasil nilai peramalan ke-t

N: Banyak data ke-n

- Perhitungan Stochastic Gradient Descent

$$\begin{aligned} & \bullet \left| \frac{47.50 - 28.14}{47.50} \right| = 0.407 \\ & \bullet \left| \frac{44.20 - 25.50}{44.20} \right| = 0.423 \\ & \bullet \left| \frac{44.11 - 22.87}{44.11} \right| = 0.481 \\ & \bullet \left| \frac{41.54 - 20.24}{41.54} \right| = 0.512 \end{aligned}$$

Total Mape:

$$\begin{aligned} \bullet \text{Mape} &= \frac{0.407 + 0.423 + 0.481 + 0.512}{4} \\ &= \frac{1.823}{4} \\ &= 0.456 \text{ Atau } 45.6\% \end{aligned}$$

- Perhitungan Support Vector Descent (SVM)

$$\begin{aligned} & \bullet \left| \frac{47.50 - 46.25}{47.50} \right| = 0.026 \\ & \bullet \left| \frac{44.20 - 45.86}{44.20} \right| = 0.037 \\ & \bullet \left| \frac{44.11 - 45.46}{44.11} \right| = 0.030 \\ & \bullet \left| \frac{41.54 - 45.09}{41.54} \right| = 0.085 \end{aligned}$$

Total Mape:

$$\begin{aligned} \bullet \text{Mape} &= \frac{0.026 + 0.037 + 0.030 + 0.085}{4} \\ &= \frac{0.177}{4} \\ &= 0.045 \text{ Atau } 4.5\% \end{aligned}$$

- Perhitungan Naural Network

$$\begin{aligned} & \bullet \left| \frac{47.50 - 7.88}{47.50} \right| = 0.834 \\ & \bullet \left| \frac{44.20 - 8.39}{44.20} \right| = 0.810 \\ & \bullet \left| \frac{44.11 - 8.91}{44.11} \right| = 0.799 \\ & \bullet \left| \frac{41.54 - 9.46}{41.54} \right| = 0.772 \end{aligned}$$

Total Mape:

$$\begin{aligned} \bullet \text{Mape} &= \frac{0.834 + 0.810 + 0.799 + 0.772}{4} \\ &= \frac{3.215}{4} \\ &= 0.804 \text{ Atau } 80.4\% \end{aligned}$$

- Hasil Perhitungan jumlah Prediksi Pada Tahun 2025 dan 2026

- **2025**

- SVM = 44.79

- SGD = 17.60

- Neural Network = 10.03

$$\text{JUMLAH} = 44.79 + 17.60 + 10.03 = 72.42$$

- **2026**

- SVM = 44.57

- SGD = 14.97

- Neural Network = 10.61

$$\text{JUMLAH} = 44.57 + 14.97 + 10.61 = 70.15$$

Dapat disimpulkan untuk hasil jumlah pada tahun 2025 dan 2026 dengan metode *SVM*, *SGD* & *Naural Network*. Adalah sebagai berikut:

Jawaban Akhir:

- Jumlah prediksi tahun 2025 = 72.42

- Jumlah prediksi tahun 2026 = 70.15

IV. KESIMPULAN

Penelitian ini menunjukkan bahwa penerapan metode data mining, khususnya algoritma Support Vector Machine (SVM), Neural Network, dan Stochastic Gradient Descent (SGD), dapat digunakan untuk memprediksi angka kemiskinan di Kabupaten Ogan Komering Ulu (OKU) dengan hasil evaluasi yang berbeda-beda tingkat akurasi. Kesimpulan ditulis dalam bentuk paragraf uraian. Hindari penggunaan *bulleted list*.

Berdasarkan hasil pengujian menggunakan metrik Mean Absolute Percentage Error (MAPE), diperoleh bahwa model dengan algoritma SVM memberikan hasil prediksi paling akurat dengan nilai MAPE sebesar **4.5%**, yang tergolong dalam kategori **sangat baik**. Sementara itu, algoritma SGD menghasilkan nilai MAPE sebesar **45.6%**, dan Neural Network sebesar **80.4%**, keduanya menunjukkan tingkat akurasi prediksi yang rendah.

Prediksi angka kemiskinan untuk tahun 2025 dan 2026 menggunakan ketiga algoritma tersebut menghasilkan jumlah prediksi total masing-masing sebesar **72.42 ribu jiwa** dan **70.15 ribu jiwa**

Adapun keunggulan metode ini adalah kemampuannya dalam mengelola dan memproses data historis dengan cepat melalui platform *Orange* yang mendukung visualisasi dan integrasi model secara efisien. Namun demikian, penelitian ini masih memiliki keterbatasan pada jumlah atribut yang digunakan, yang hanya mencakup variabel tahun dan angka kemiskinan.

Untuk pengembangan selanjutnya, disarankan agar penelitian melibatkan lebih banyak variabel independen yang relevan seperti tingkat pengangguran, pendidikan, dan kesehatan. Selain itu, penggunaan algoritma ensemble atau hybrid dapat menjadi alternatif untuk meningkatkan akurasi prediksi dalam penelitian sejenis.

UCAPAN TERIMA KASIH

Penulis mengucapkan terima kasih kepada pihak-pihak terkait yang telah memberi dukungan terhadap penelitian ini. Sehingga Penulis dapat menyelesaikan jurnal ini [15].

REFERENSI

- [1] Khalim, K. A., Hayati, U., & Bahtiar, A. (2023). Perbandingan Prediksi Penyakit Hipertensi Menggunakan Metode Random Forest Dan Naïve Bayes. *JATI (Jurnal Mahasiswa Teknik Informatika)*, 7(1), 498-504.
- [2] Kusuma, F., Ahsan, M., & Syahminan, S. (2021). Prediksi Jumlah Penduduk Miskin Indonesia menggunakan Metode Single Moving Average dan Double Moving Average. *Jurnal Informatika dan Rekayasa Perangkat Lunak*, 3(2), 105-109.
- [3] Ridwan, M., Suyono, H., & Sarosa, M. (2013). Penerapan data mining untuk evaluasi kinerja akademik mahasiswa menggunakan algoritma naive bayes classifier. *jurnal EECCIS*, 7(1), 59-64.
- [4] Rosida, P. L., Nurmalasari, M., Hosizah, H., & Krismawati, D. (2024). Implementasi Decision Tree untuk Prediksi Kelahiran Bayi Prematur. *Jurnal Manajemen Informatika (JAMIKA)*, 14(2), 178-186.
- [5] Arifin, J. (2020). Budaya kemiskinan dalam penanggulangan kemiskinan di Indonesia. *Sosio Informa*, 6(2), 114-132.
- [6] Tisniwati, B. (2012). Analisis faktor-faktor yang mempengaruhi tingkat kemiskinan di Indonesia. *Jurnal Ekonomi Pembangunan*, 10(1), 33-46.
- [7] Sinaga, L. M., & Sipayung, S. P. (2024). Analisis dan Prediksi Persentase Angka Kemiskinan di Indonesia menggunakan Metode Regresi Linier Berganda. *KAKIFIKOM (Kumpulan Artikel Karya Ilmiah Fakultas Ilmu Komputer)*, 116-125.
- [8] Yallah, H., Darwis, M., & Henrowati, R. (2024). PENERAPAN ALGORITMA ID3 MELALUI APLIKASI ORANGE UNTUK PREDIKSI AKURASI AKREDITASI SEKOLAH DASAR DI DEPOK. *Jurnal Manajemen Informatika dan Sistem Informasi*, 7(2), 131-141.
- [9] Haryati, S., Sudarsono, A., & Suryana, E. (2015). implementasi data mining untuk memprediksi masa studi mahasiswa menggunakan algoritma c4. 5 (studi kasus: universitas dehasen bengkulu). *Jurnal Media Infotama*, 11(2).
- [10] Aprilia, K., & Sembiring, F. (2021, September). Analisis Garis Kemiskinan Makanan Menggunakan Metode Algoritma K-Means Clustering. In *Prosiding Seminar Nasional Sistem Informasi Dan Manajemen Informatika Universitas Nusa Putra* (Vol. 1, pp. 1-10).
- [11] Fauzi, R. N., Febriani, R. K., & Desmawan, D. (2022). Pengaruh Laju Pertumbuhan Penduduk Terhadap Kemiskinan Di Indonesia. *Jurnal Ekonomi, Bisnis Dan Manajemen*, 1(1), 118-122.
- [12] Ginting, R., & Humaira, C. (2021, October). Penerapan Data Mining: Prediksi Penjualan Mobil Toyota Menggunakan Artificial Neural Network pada Software Orange. In *Talenta Conference Series: Energy and Engineering (EE)* (Vol. 4, No. 1).
- [13] Muharrom, M. (2023). Analisis Komparasi Algoritma Data Mining Naive Bayes, K-Nearest Neighbors dan Regresi Linier Dalam Prediksi Harga Emas. *Bulletin of Information Technology (BIT)*, 4(4), 430-438.
- [14] Prasetya, R. (2020). Penerapan Teknik Data Mining Dengan Algoritma Classification Tree Untuk Prediksi Hujan. *Jurnal Widya Climago*, 2(2).
- [15] Soedjarwanto, N., & Prayoga, J. (2025). ANALISIS PERBANDINGAN TEMPERATURE WINDING PADA MOTOR SEPARATOR AREA RAW MILL DAN CEMENT MILL DI PT SEMEN BATURAJA Tbk. *Jurnal Informatika dan Teknik Elektro Terapan*, 13(1).