

# Penerapan *Data Mining* Untuk Memprediksi Jumlah Total Produksi Hcl Pada Perusahaan PT.Lontar Papyrus Menggunakan Algoritma Regresi Linier Berganda

M. Al Hafizh Laksamana<sup>1</sup>, Amroni<sup>2</sup>, Afrizal Nehemia Toscani<sup>3</sup>

Teknik Informatika, Fakultas Ilmu Komputer, Universitas Dinamika Bangsa  
Jalan Jendral Sudirman Thehok, Jambi, Indonesia  
Email : alhafizh1701@gmail.com, amroni69@yahoo.com, afrizal@stikom-db.ac.id

## ABSTRACT

PT. Lontar Papyrus Pulp and Paper Industry (PT. LPPPI) is one of the companies in the APP Group (Asia Pulp and Paper) which was established in the Year of this pulp and tissue mill located in Tebing Tinggi Village, Tungkal Ulu District, Tanjung Jabung Barat District, Province Jambi HCl is an aquatic solution from hydrogen chloride gas. He is a strong acid. This compound is also widely used in industry. One of them is in the Pulp manufacturing industry at PT.Lontar Papyrus. The use of HCl in the manufacture of Pulp alone serves as an accelerating reaction. Based on this, it is necessary to study the concentration of the catalyst and the right cooking time to produce pulp with the best and perfect chemical properties. At PT. Lontar Papyrus the amount of HCl production still hasn't made predictions in production in the next month so that production is still inefficient. The problem that the writer found in this case is the amount of HCl production in PT. Paparrus papyrus sometimes excess so that the amount of production does not match what was wanted before and becomes less efficient. Of the problems that the authors found in the Company PT. Lontar Papyrus, the author is interested in conducting research as outlined in the form of scientific writing entitled "The Application of Data Mining to Predict the Total Amount of Hcl Production in PT.Lontar Papyrus Company Using Multiple Linear Regression Algorithms.

*Keywords:* implementation, data mining, multiple linear regression

## ABSTRAK

PT. Lontar Papyrus Pulp and Paper Industry (PT. LPPPI) adalah salah satu perusahaan di dalam Group APP (Asia Pulp and Paper) yang didirikan pada Tahun Pabrik pulp dan Tissue ini berlokasi di Desa Tebing Tinggi, Kecamatan Tungkal Ulu, Kabupaten Tanjung Jabung Barat, Propinsi Jambi HCl merupakan larutan akuatik dari gas hidrogen klorida. Ia adalah asam kuat. Senyawa ini juga digunakan secara luas dalam industri. Salah satunya dalam industri pembuatan Pulp di PT.Lontar Papyrus. Penggunaan HCl dalam pembuatan Pulp tersendiri berfungsi sebagai mempercepat reaksi. Berdasarkan hal tersebut, maka perlu dipelajari konsentrasi katalis dan lama pemasakan yang tepat untuk menghasilkan pulp dengan sifat kimia terbaik dan juga sempurna. Pada PT. Lontar Papyrus jumlah produksi HCl masih belum melakukan prediksi dalam produksi dalam bulan depannya sehingga produksi masih belum efisien. Masalah yang penulis temukan dalam hal ini adalah jumlah produksi HCl di PT.Lontar Papyrus terkadang berlebih sehingga jumlah produksi tidak sesuai dengan yang di inginkan sebelumnya dan menjadi kurang efisien. Dari masalah yang penulis temukan di Perusahaan PT. Lontar Papyrus, penulis tertarik melakukan penelitian yang dituangkan dalam bentuk penulisan ilmiah yang berjudul "Penerapan *Data Mining* Untuk Memprediksi Jumlah Total Produksi Hcl Pada Perusahaan PT.Lontar Papyrus Menggunakan Algoritma Regresi Linier Berganda.

*Kata kunci:* penerapan, data mining, regresi linier berganda

## 1. PENDAHULUAN

Dengan pesatnya perkembangan teknologi informasi, hampir semua perusahaan menggunakan teknologi informasi untuk mengolah data, termasuk memproses, mendapatkan, menyusun, menyimpan, memanipulasi data dalam berbagai cara untuk menghasilkan informasi yang berkualitas, yaitu informasi yang relevan, akurat dan tepat waktu sehingga informasi yang dihasilkan dapat membantu dalam pengambilan keputusan yang tepat, akurat, efisien.

*Data mining* adalah proses menemukan hubungan dalam data yang tidak diketahui oleh pengguna dan menyajikannya dengan cara yang dapat dipahami sehingga hubungan tersebut dapat menjadi dasar pengambilan keputusan [1].

HCl merupakan larutan akuatik dari gas hidrogen klorida. Ia adalah asam kuat. Senyawa ini juga digunakan secara luas dalam industri. Salah satunya dalam industri pembuatan Pulp di PT.Lontar Papyrus. Penggunaan HCl dalam pembuatan Pulp tersendiri berfungsi sebagai mempercepat reaksi. Berdasarkan hal tersebut, maka perlu dipelajari konsentrasi katalis dan lama pemasakan yang tepat untuk menghasilkan pulp dengan sifat kimia terbaik dan juga sempurna. Pada PT. Lontar Papyrus jumlah produksi HCl masih belum melakukan prediksi dalam produksi dalam bulan depannya sehingga produksi masih belum efisien.

Penelitian ini memanfaatkan data produksi HCl di PT. Lontar Papyrus dengan menerapkan salah satu teknik dari data mining yaitu *Regresi Linier*. Hal ini dilakukan dengan harapan supaya dapat memprediksi hasil produksi HCL di PT. Lontar Papyrus agar dapat lebih efisien dalam memproduksi HCl itu sendiri sehingga tidak ada produksi yang berlebihan mau pun kekurangan.

## 2. TINJAUAN PUSTAKA

### 2.1. Data Mining

Data mining merupakan disiplin ilmu yang mempelajari metode untuk mengekstrak pengetahuan atau menemukan pola dari suatu data., Heni Sulastrri (2017:299) [2].

Menurut Yuli Mardi (2016:2) mengatakan : *Data mining* adalah proses mencari pola atau informasi menarik dalam data terpilih dengan menggunakan teknik atau metode tertentu [3].

Serangkaian proses mendapatkan pengetahuan atau pola dari kumpulan data disebut dengan data mining, Khafiizh Hastuti (2012:242) [4].

### 2.2. Regresi Linier

Deny Kurniawan (2008:3) Regresi linier adalah metode statistika yang digunakan untuk membentuk model hubungan antara variabel terikat (dependen; respon; Y) dengan satu atau lebih variabel bebas (independen, prediktor, X) [5]. Apabila banyaknya variabel bebas hanya ada satu, disebut sebagai regresi linier sederhana, sedangkan apabila terdapat lebih dari 1 variabel bebas, disebut sebagai regresi linier berganda.

Analisis regresi setidaknya-tidaknya memiliki 3 kegunaan, yaitu untuk tujuan deskripsi dari fenomena data atau kasus yang sedang diteliti, untuk tujuan kontrol, serta untuk tujuan prediksi. Regresi mampu mendeskripsikan fenomena data melalui terbentuknya suatu model hubungan yang bersifat numerik. Regresi juga dapat digunakan untuk melakukan pengendalian (kontrol) terhadap suatu kasus atau hal-hal yang sedang diamati melalui penggunaan model regresi yang diperoleh. Selain itu, model regresi juga dapat dimanfaatkan untuk melakukan prediksi untuk variabel terikat. Namun yang perlu diingat, prediksi di dalam konsep regresi hanya boleh dilakukan di dalam rentang data dari variabel-variabel bebas yang digunakan untuk membentuk model regresi tersebut. Misal, suatu model regresi diperoleh dengan mempergunakan data variabel bebas yang memiliki rentang antara 5 s.d. 25, maka prediksi hanya boleh dilakukan bila suatu nilai yang digunakan sebagai input untuk variabel X berada di dalam rentang tersebut. Konsep ini disebut sebagai interpolasi.

Data untuk variabel independen X pada regresi linier bisa merupakan data pengamatan yang tidak ditetapkan sebelumnya oleh peneliti (observational data) maupun data yang telah ditetapkan (dikontrol) oleh peneliti sebelumnya (*experimental or fixed data*). Perbedaannya adalah bahwa dengan menggunakan fixed data, informasi yang diperoleh lebih kuat dalam menjelaskan hubungan sebab akibat antara variabel X dan variabel Y. Sedangkan, pada observational data, informasi yang diperoleh belum tentu merupakan hubungan sebab-akibat. Untuk fixed data, peneliti sebelumnya telah memiliki beberapa nilai variabel X yang ingin diteliti. Sedangkan, pada observational data, variabel X yang diamati bisa berapa saja, tergantung keadaan di lapangan. Biasanya, fixed data diperoleh dari percobaan laboratorium, dan observational data diperoleh dengan menggunakan kuesioner.

### 2.2.1. Regresi Linier Berganda

Amrin (2016:75) Analisis yang memiliki variabel bebas lebih dari satu disebut analisis regresi linier berganda. Teknik regresi linier berganda digunakan untuk mengetahui ada tidaknya pengaruh signifikan dua atau lebih variabel bebas ( $X_1, X_2, X_3, \dots, X_k$ ) terhadap variabel terikat ( $Y$ ) [6].

Jamner R. Lawendatu (2014:67) Analisis yang memiliki variabel bebas lebih dari satu disebut analisis regresi berganda. Analisis regresi linier berganda memberikan kemudahan bagi pengguna untuk memasukkan lebih dari satu variabel bebas hingga  $k$  dimana banyaknya  $k$  kurang dari jumlah observasi ( $n$ ) [7].

Metode Regresi Linear Berganda merupakan pendekatan yang paling tepat untuk meramalkan angka kelahiran dimasa mendatang. Analisis yang memiliki variabel bebas lebih dari satu disebut analisis regresi linier berganda. Nike Syafitri (2018:31) [8].

### 2.3. Tahapan Proses Dalam Data Mining

*Knowledge Discovery In Database (KDD)* merupakan metode untuk memperoleh pengetahuan dari database yang ada. Dalam database terdapat tabel - tabel yang saling berhubungan / berelasi. Hasil pengetahuan yang diperoleh dalam proses tersebut dapat digunakan sebagai basis pengetahuan (knowledge base) untuk keperluan pengambilan keputusan.

Istilah *Knowledge Discovery in Database (KDD)* dan data mining seringkali digunakan secara bergantian untuk menjelaskan proses penggalian informasi tersembunyi dalam suatu basis data yang besar. Sebenarnya kedua istilah tersebut memiliki konsep yang berbeda, tetapi berkaitan satu sama lain, dan salah satu tahapan dalam keseluruhan proses KDD adalah data mining. Proses KDD secara garis besar dapat dijelaskan sebagai berikut (Yuli Mardi, 2016:213): [9].

#### 2.3.1. Data Selection

Pemilihan (seleksi) data dari sekumpulan data operasional perlu dilakukan sebelum tahap penggalian informasi dalam *Knowledge Discovery in Database(KDD)* dimulai. Data hasil seleksi yang akan digunakan untuk proses data mining, disimpan dalam suatu berkas terpisah dari basis data operasional.

#### 2.3.2. Preprocessing Cleaning

Sebelum proses data mining dapat dilaksanakan, perlu dilakukan proses cleaning pada data yang menjadi fokus *Knowledge Discovery in Database (KDD)*. Proses cleaning mencakup antara lain membuang duplikasi data, memeriksa data yang inkonsisten, dan memperbaiki kesalahan pada data, seperti kesalahan cetak. Juga dilakukan proses enrichment, yaitu proses “memperkaya” data yang sudah ada dengan data atau informasi lain yang relevan dan diperlukan untuk *Knowledge Discovery in Database (KDD)*, seperti data atau informasi eksternal lainnya yang diperlukan.

#### 2.3.3. Transformation

Coding adalah proses transformasi pada data yang telah dipilih, sehingga data tersebut sesuai untuk proses data mining. Proses coding dalam *Knowledge Discovery in Database (KDD)* merupakan proses kreatif dan sangat tergantung pada jenis atau pola informasi yang akan dicari dalam basis data.

#### 2.3.4. Data mining

Data mining adalah proses mencari pola atau informasi menarik dalam data terpilih dengan menggunakan teknik atau metode tertentu. Teknik-teknik, metode-metode, atau algoritma dalam data mining sangat bervariasi. Pemilihan metode atau algoritma yang tepat sangat bergantung pada tujuan dan proses *Knowledge Discovery in Database (KDD)* secara keseluruhan.

#### 2.3.5. Interpretation/Evaluation

Pola informasi yang dihasilkan dari proses data mining perlu ditampilkan dalam bentuk yang mudah dimengerti oleh pihak yang berkepentingan. Tahap ini merupakan bagian dari proses *Knowledge Discovery in Database (KDD)* yang disebut interpretation. Tahap ini mencakup pemeriksaan apakah pola atau informasi yang ditemukan bertentangan dengan fakta atau hipotesis yang ada sebelumnya.

#### 2.4. *Pengelompokan Data Mining*

Menurut Larose dalam bukunya yang berjudul "Discovering Knowledge in Data: An Introduction to Data Mining", data mining dibagi menjadi beberapa kelompok berdasarkan tugas/pekerjaan yang dapat dilakukan, yaitu (Larose dalam jurnal Goldie Gunadi dan Dana Indra, 2012: 120) : [10].

##### 2.4.1. *Deskripsi*

Terkadang penelitian dan analisis secara sederhana ingin mencoba mencari cara untuk menggambarkan pola dan kecenderungan yang terdapat dalam data. Sebagai contoh, petugas pengumpul suara mungkin tidak menemukan keterangan atau fakta bahwa siapa yang tidak cukup profesional atau sedikit didukung dalam pemilihan presiden. Deskripsi dari pola dan kecenderungan sering memberikan kemungkinan penjelasan untuk suatu pola atau kecenderungan.

##### 2.4.2. *Estimasi*

Estimasi hampir sama dengan klasifikasi, kecuali variable target estimasi lebih ke arah numeric dari pada ke arah kategori. Model dibangun dengan *record* lengkap menyediakan nilai dari variable target sebagai nilai prediksi. Selanjutnya, pada peninjauan berikutnya estimasi nilai dari variable target dibuat berdasarkan variable prediksi.

##### 2.4.3. *Prediksi*

Prediksi hampir sama dengan klasifikasi dan estimasi, kecuali bahwa dalam prediksi nilai dari hasil akan datang. Beberapa metode dan teknik yang digunakan dalam klasifikasi dan estimasi dapat pula digunakan (untuk keadaan yang tepat) untuk prediksi.

##### 2.4.4. *Klasifikasi*

Dalam klasifikasi, terdapat target variable kategori. Sebagai contoh, penggolongan pendapat dapat dipisahkan dalam tiga kategori, yaitu pendapatan tinggi, pendapatan sedang, dan pendapatan rendah.

##### 2.4.5. *Pengklusteran*

Pengklusteran merupakan pengelompokan *record*, pengamatan, atau memperhatikan dan membentuk kelas objek-objek yang memiliki kemiripan. Kluster adalah kumpulan *record* yang memiliki kemiripan satu dengan yang lainnya dan memiliki ketidakmiripan dengan *record-record* dalam kluster lain. Pengklusteran berbeda dengan klasifikasi yaitu tidak adanya variable target dalam pengklusteran. Pengklusteran tidak mencoba untuk melakukan klasifikasi, mengestimasi, atau memprediksi nilai dari variable target. Akan tetapi, algoritma pengklusteran mencoba untuk melakukan pembagian terhadap keseluruhan data menjadi kelompok-kelompok yang memiliki kemiripan *record* dalam satu kelompok akan bernilai maksimal, sedangkan kemiripan dengan *record* dalam kelompok lain akan bernilai minimal.

##### 2.4.6. *Asosiasi*

Tugas asosiasi dalam *data mining* adalah menemukan atribut yang muncul dalam satu waktu. Dalam dunia bisnis lebih umum disebut analisis keranjang belanja (*market basket analysis*).

#### 2.5. *Tools Data Mining*

Berikut ini penulis akan sedikit menjelaskan beberapa aplikasi/tools yang umum digunakan dalam data mining :

##### 2.5.1. *Spss*

SPSS (Statistical Program for Social Science) merupakan paket program yang berguna untuk menganalisis data statistik. SPSS dapat digunakan untuk hampir seluruh file data dan sekaligus membuat laporan dalam bentuk tabulasi, grafik, dan plot untuk berbagai distribusi maupun statistic deskriptif. [11]. Semula SPSS hanya digunakan untuk ilmu social saja, tapi perkembangan berikutnya digunakan untuk berbagai disiplin ilmu sehingga kepanjangannya berubah menjadi "Statistical Product and Service Solution" (Nisfiannoor, Muhammad, Pendekatan Statistika Modern Untuk Ilmu Social [12]. SPSS digunakan oleh peneliti pasar, peneliti kesehatan, perusahaan survei, pemerintah, peneliti pendidikan, organisasi pemasaran, dan sebagainya. Selain analisis statistika, manajemen data (seleksi kasus, penajaman file, pembuatan data turunan) dan dokumentasi data (kamus metadata ikut dimasukkan bersama data) juga merupakan fitur-fitur dari software dasar SPSS.

### 2.6. Penelitian Sejenis

Penelitian yang dilakukan oleh Yuliana pada tahun 2014 dengan judul Penerapan Model Regresi Linear Robust Dengan Estimasi M pada Data Nilai Kalkulus II Mahasiswa Universitas Widya Dharma Klaten. Penelitian ini menghasilkan sebuah Melalui regresi linear robust dengan estimasi M diperoleh suatu estimasi parameter regresi yang konvergen tanpa harus membuang pengamatan *outliernya*. Hal ini berarti regresi linear robust dengan estimasi M dapat digunakan untuk mengatasi suatu data yang mengandung pengamatan *outlier*. Dari model regresi robust yang telah didapat tersebut diperoleh suatu model regresi robust persamaan  $Y = 0.177X_1 + 0.655X_2 - 0.947$ . dari model regresi linear robust ini dapat digunakan untuk memprediksi suatu nilai kalkulus ii secara tepat..

Selanjutnya pada tahun 2019, Arief Kurniadi juga melakukan penelitian dengan judul Penerapan Metode Regresi Linier Untuk Memprediksi Kebiasaan Pelanggan Studi Kasus: PT.Mensa Binasukses. Hasil penelitiannya. Hubungan antara pesanan, penjualan, dan saldo mempunyai hubungan yang kuat positif dan searah, dari hasil analisis regresi linear berganda yaitu  $y = 18.512 - 0.573129 + 0.710729$ . Setelah dilakukan analisis linear pada tools WEKA menggunakan *10-Fold cross Validation* , maka didapatlah akurasi tertinggi yaitu pesanan terbanyak adalah 7038 dengan rata-rata 635,367 nama obat FINPRO BOX 30 TABLET. dari model regresi linear ini dapat digunakan untuk memprediksi suatu nilai kalkulus ii secara tepat.

Penelitian lainnya dilakukan oleh Farizal, Amar Rachman, dan Hadi Al Rasyid pada tahun 2014 dengan judul Model Peramalan Konsumsi Bahan Bakar Jenis Premium di Indonesia dengan Regresi Linier Berganda. Penelitian ini memberikan alternative teknik estimasi konsumsi premium yang lebih baik dari yang ada (digunakan) selama ini yang ditunjukkan dengan tingkat kesalahan estimasi yang lebih rendah.

## 3. METODOLOGI PENELITIAN

Penelitian ini dilaksanakan melalui beberapa tahapan sebagai berikut:

### 1. Studi Literatur

Pada tahap ini dilakukan pengumpulan data dengan cara membaca buku melalui literatur dan buku lain yang bersifat ilmiah dan juga internet yang ada hubungannya dengan penelitian untuk melengkapi pembendaharaan konsep dan teori, sehingga memiliki landasan dan keilmuan yang baik dan sesuai.

### 2. Identifikasi Masalah

Masalah yang diidentifikasi dalam penelitian ini adalah bagaimana penerapan metode *Regresi Linier* dalam memprediksi produksi pada perusahaan PT.Lontar Papyrus.

### 3. Pengumpulan Data

Dalam pengumpulan data, penulis mendapatkan data Perusahaan PT Lontar Papyrus.

### 4. Pengelompokan Item Produksi

Pada tahap ini dilakukan pengelompokkan item produksi atau bahan-bahan apa saja yang memiliki ikatan, agar data tersebut dapat dikelola informasinya menggunakan Microsoft office excel 2016 sehingga proses data produksi dapat dilakukan *tools SPSS* menghasilkan *output* berupa sebuah informasi produksi

### 5. Seleksi Atribut

Pada tahap ini pihak PT. Lontar Papyrus dapat menseleksi bahan-bahan produksi, manakah bahan produksi yang sering dipakai dan yang kurang untuk dipakai, agar PT. Lontar Papyrus dapat menyediakan bahan produksi tidak terlalu berlebihan

### 6. Hasil Analisis

Pada tahap ini data produksi yang dilakukan proses *Data Mining* akan menghasilkan nilai nantinya akan ditampilkan dalam bentuk representasi grafik, sehingga informasi yang menentukan bagaimana produksi yang akan dilakukan pada bulan berikutnya.

## 7. Pembuatan Laporan

Setelah semua tahapan penelitian dilakukan, maka akan dibuat laporan sebagai dokumentasi penelitian agar dapat dimanfaatkan pada waktu yang akan datang. Baik oleh peneliti sendiri maupun peneliti lainnya.

## 4. ANALISIS DAN HASIL

### 4.5. Analisis Perhitungan

Analisis perhitungan ini menjelaskan langkah-langkah perhitungan prediksi jumlah produksi HCL yang menggunakan Algoritma Regresi Linier Berganda dengan variabel dependen : Jumlah Produksi HCL dan variabel independen : Pure Water perbulan dan CL2(weak) perbulan.

### 4.6. Hasil Perhitungan

#### 4.2.1 Multiple R

Atau *R* Majemuk yaitu untuk mengukur suatu tingkat keeratan hubungan *linear* antara variabel dependen (*Y*) yaitu jumlah produksi HCL dengan seluruh variabel independen yaitu Pure Water (*X*<sub>1</sub>) dan CL2 Weak (*X*<sub>2</sub>) secara bersama-sama maka dilakukan perhitungan sebagai berikut :

$$\text{Multiple } R = (R \text{ Square})^{0.5}$$

Dimana nilai *R Square* yaitu sebesar 0,951120875. Maka *Multiple R* = (0,823776069)<sup>0.5</sup> = 0,907621104

#### 4.2.2 R Square (*R*<sup>2</sup>)

*R Square* atau koefisien determinasi berganda. Berfungsi untuk mengukur kebaikan suai (*goodness of fit*) dari persamaan regresi, yaitu memberikan proporsi atau persentase variasi total dalam variabel dependen yang dijelaskan oleh variabel independen. Adapun perhitungan dari *R Square* yaitu :

$$R \text{ Square} = 1 - \left( \frac{SS_{res}}{SS_{total}} \right)$$

Dimana nilai *SSres* = 1236799,028 dan *SStotal* = 7018337,517. Nilai *SSres* dan *SStotal* dapat dilihat pada tabel ANOVA kolom *SS (Sum of Square)*

$$\text{Maka } R \text{ Square} = 1 - \left( \frac{1236799,028}{7018337,517} \right) = 0,823776069$$

#### 4.2.3 Adjusted R Square

Digunakan untuk membandingkan dua *R*<sup>2</sup> dari dua model (*X* dan *Y*) dan memperhitungkan banyaknya variabel independen yang ada dalam model, karena suatu sifat penting *R*<sup>2</sup> adalah nilainya merupakan fungsi yang tidak pernah menurun dari banyaknya variabel independen yang

ada dalam model. Maka *Adjusted R Square* =  $1 - \frac{\left( \frac{1236799,028}{9} \right)}{\left( \frac{7018337,517}{11} \right)} = 0,784615195$

#### 4.2.4 Standard Error

Merupakan *standar deviasi error* dari estimasi variabel dependen (Jumlah Produksi HCL), yaitu 370,7048883

#### 4.2.5 Observations

Jumlah data yang peneliti uji dalam perhitungan ini. Adapun hasil nya dapat dilihat pada tabel 1 :

**Tabel 1 Summary Output**

<i>Regression Statistics</i>	
Multiple R	0,907621104
R Square	0,823776069
Adjusted R Square	0,784615195
Standard Error	370,7048883
Observations	12

#### 4.3 Hasil Prediksi

Adapun untuk mengetahui *output* dari *regresi linear* berganda ini dapat dilihat pada tabel 2 :

**Tabel 2 Hasil Prediksi HCL Tahun 2020 (TON)**

<b>Bulan</b>	<b>Prediksi HCL (Y)</b>
Januari	7631,855111
Februari	8555,651072
Maret	8369,339359
April	7178,800357
Mei	8014,644625
Juni	9088,4535
Juli	8347,410875
Agustus	8600,02836
September	7586,349567
Oktober	6449,634975
November	8240,483688
Desember	8505,067746

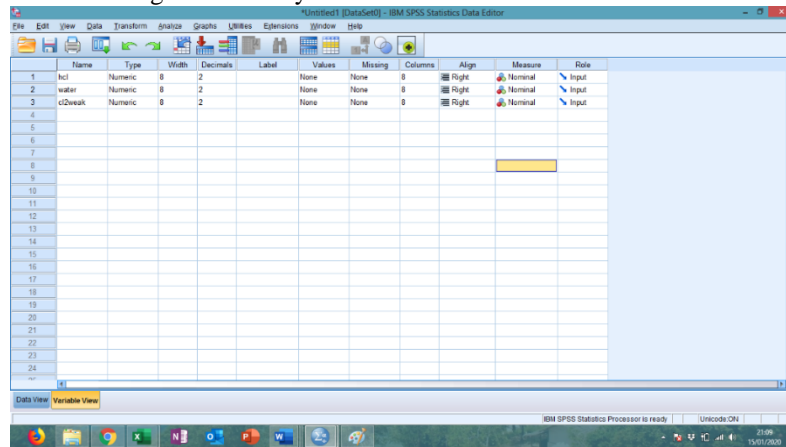
Terlihat diatas merupakan hasil prediksi Jumlah HCL pada tahun 2020 yang di dapat dari data dari 3 tahun yang lalu 2017,2018 2019 yang telah di rata-ratakan.

#### 4.4 Hasil Analisis

Hasil analisis merupakan hasil yang telah didapatkan sebelumnya yang di implementasikan di dalam sebuah aplikasi bernama spss (Statistical Package for the Social Sciences)

### 1. Tahapan Proses Data

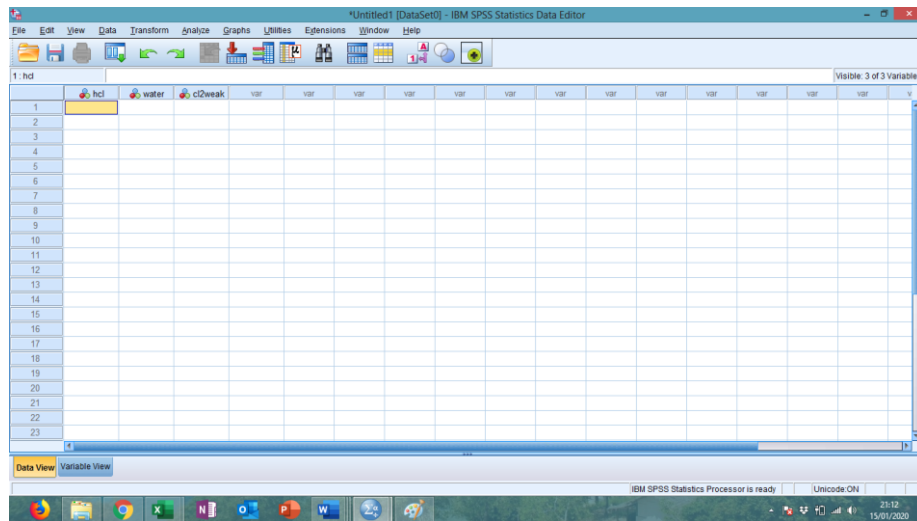
Tahapan proses ini dilakukan dengan memberikan properti pada kolom *Name* di variabel *view* dan pada kolom Label diberi nama sesuai keterangan variabelnya



Gambar 1. Memberikan Properti di Variabel View

### 2. Tabulasi data

Pada tahap tabulasi data ini, semua data dimasukkan ke dalam data *view* yang telah dibuat di variabel *view*.

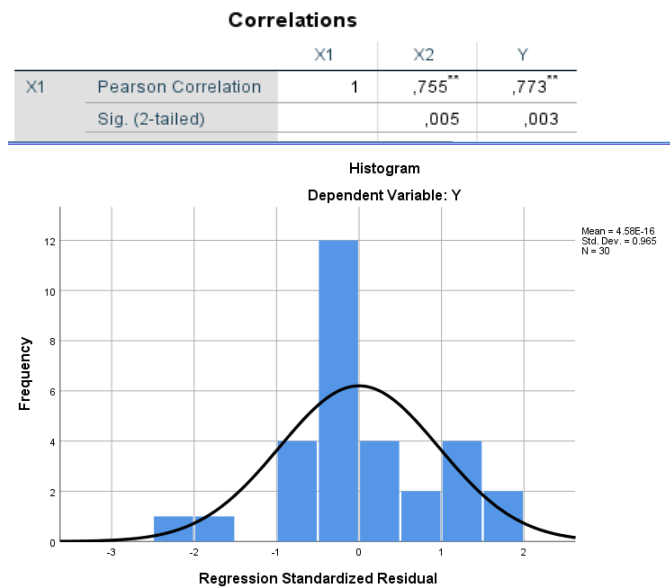


Gambar 2. Tabulasi Data

### 3. Uji Validitas

Uji validitas dilakukan untuk mengukur sah atau tidaknya indikator atau instrument kuesioner dari masing-masing variabel. Pengujian dilakukan dengan membandingkan  $r$  hitung dan  $r$  tabel. Nilai  $r$  hitung merupakan hasil korelasi jawaban responden pada masing-masing pertanyaan dengan total jawaban untuk masing-masing variabel setiap item pertanyaan/instrument disebut valid apabila  $r$  hitung lebih besar dibandingkan  $r$  tabel.





Gambar 3. Uji Validitas

#### 4. Membaca Output

Output yang dikeluarkan ketika melakukan pengujian regresi berganda terdiri dari beberapa tabel yaitu tabel variable entered/remove, model summary, anova, dan coefficients.

**Coefficients<sup>a</sup>**

Model		Unstandardized Coefficients		Standardized Coefficients	t	Sig.	Correlations			Collinearity Statistics	
		B	Std. Error	Beta			Zero-order	Partial	Part	Tolerance	VIF
1	(Constant)	663,386	1220,335		,544	,600					
	X1	,325	,308	,225	1,055	,319	,773	,332	,148	,431	2,322
	X2	6,496	1,908	,726	3,404	,008	,896	,750	,476	,431	2,322

a. Dependent Variable: Y

Gambar 4. Coefficients

#### 5. Anova

Uji F diuji dengan menggunakan uji keseluruhan, dengan taraf signifikan 5% (0,05). Jika F hitung < F tabel, maka Ho diterima Jika F hitung  $\geq$  F tabel, maka H1 ditolak

Gambar 5. Anova.

#### 6. Nilai t dengan kurva

Dengan melihat perbandingan dengan kurva bisa mengetahui arah pengaruh dari variabel X yang berarti pengaruh positif atau negatif

Gambar 6. Nilai t dengan Kurva

**ANOVA<sup>a</sup>**

Model		Sum of Squares	df	Mean Square	F	Sig.
1	Regression	5781538,490	2	2890769,245	21,036	,000 <sup>b</sup>
	Residual	1236799,029	9	137422,114		
	Total	7018337,519	11			

a. Dependent Variable: Y  
b. Predictors: (Constant), X2, X1

### 7. Uji Reliabilitas

Uji reliabilitas digunakan untuk mengetahui apakah indikator atau kuesioner yang digunakan dapat dipercaya sebagai alat ukur variabel.

Reliability Statistics				
Cronbach's Alpha	N of Items			
,384	2			

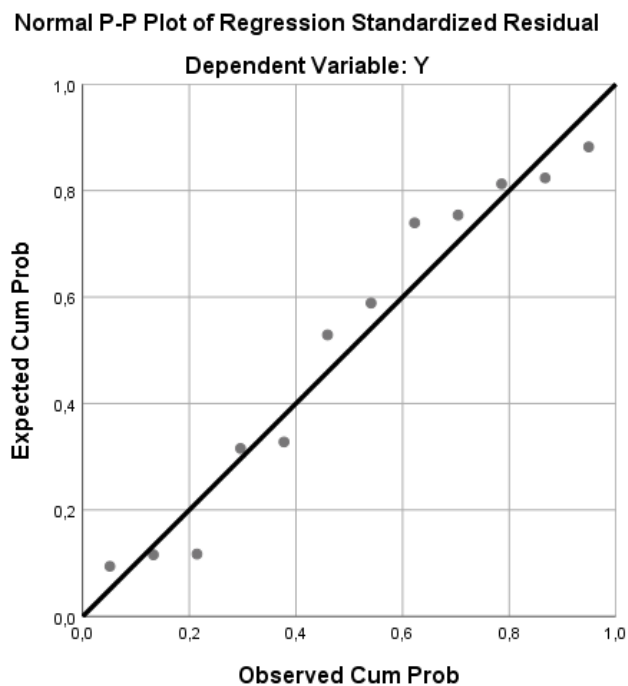
  

Item-Total Statistics				
	Scale Mean if Item Deleted	Scale Variance if Item Deleted	Corrected Item-Total Correlation	Cronbach's Alpha if Item Deleted
PureWater	841,58	7965,888	,755	.
CL2Weak	5904,00	306074,783	,755	.

Gambar 7. Uji Reliabilitas

### 8. Normalitas probability plot

Model *Regresi* dikatakan berdistribusi normal jika data *ploting* (titik-titik) yang menggambarkan data sesungguhnya mengikuti garis diagonal. Hasil *Probability Plot*



Gambar 8. Normal Probability Plot

## 5. PENUTUP

### 5.1. Kesimpulan

Setelah melalui tahap analisa dan pengujian pada perangkat lunak sistem prediksi hasil produksi menggunakan metode regresi linier berganda maka dapat diambil beberapa kesimpulan sebagai berikut:

1. Persamaan regresi yang diperoleh dari hasil analisis regresi linier berganda dari manual maupun dari aplikasi SPSS adalah sebagai berikut :

$$Y = 663,386017 + (0,32464744)X_1 + (6,49637061)X_2$$

2. Data produksi yang dianalisa pada PT Lontar Papyrus menggunakan algoritma regresi linier berganda yang akan memprediksi produksi HCL tahun 2020, dalam penelitian ini adalah menggunakan data acuan produksi HCL dalam kurun waktu 3(tiga) tahun, yaitu pada tahun 2017, 2018 dan tahun 2019 yang telah di rata-ratakan. Di dapatkan tingkat hasil akurasi dari hasil analisis ini senilai 78% akurasi. Setelah di lakukan analisis, dapat di simpulkan bahwa variabel CL2 Strong, tidak cukup tinggi mempengaruhi dalam jumlah produksi HCL maka dari itu di ganti dengan variabel Pure Water yang Lebih koefisien.
3. Hubungan antara pure water, CL2, dan produksi HCL mempunyai hubungan yang kuat positif dan searah.

### 5.2. Saran

Dari kesimpulan yang telah diuraikan diatas, maka peneliti memberikan saran saran sebagai berikut :

1. Analisa dengan menggunakan nilai batasan frekuensi yang lebih banyak untuk mengoptimalkan hasil analisis.
2. Penelitian yang akan datang bisa dikembangkan dengan menambahkan beberapa variable independen supaya diperoleh hasil yang lebih baik dari penelitian ini.
3. Untuk pengembangan selanjutnya , diharapkan pengembang menggunakan variabel independen yang memiliki tingkat pengaruh lebih tinggi terhadap variabel dependen (produksi HCL).

## DAFTAR PUSTAKA

- [1] Amrin, 2016. Data Mining Dengan Regresi Linier Berganda Untuk Peramalan Tingkat Inflasi. Jurnal Techno Nusa Mandiri Vol. XIII, No. 1 Maret 2016. Pp.74-79.
- [2] Brilian Rahmat C.T.I., Agum Agidatama Gafar, Nurul Fajriani, Umar Ramdani, Fitria Rihin Uyun, Yuwanda Purnamasari P dan Natalis Ransi7, 2017. Implementasi K-Means Clustering Pada Rapidminer Untuk Analisis Daerah Rawan Kecelakaan. Seminar Nasional Riset Kuantitatif Terapan. Pp-58-62
- [3] Deny Kurniawan, 2008. Regresi Linier (Linear Regression). Vienna, Austria. ISBN 3-900051-07-0.Pp.1-14
- [4] Goldie Gunadi, 2012. Penerapan Metode Data Mining Market Basket Analysis Terhadap Data Penjualan Produk Buku Dengan Menggunakan Algoritma Apriori Dan Frequent Pattern Growth (Fp-Growth) : Studi Kasus Percetakan Pt. Gramedia. Jurnal TELEMATIKA MKOM Vol.4 No.1. Pp. 118-132
- [5] Heni Sulastri, 2017. Penerapan Data Mining Dalam Pengelompokan Penderita Thalassaemia. Jurnal Nasional Teknologi Dan Sistem Informasi - Vol. 03 No. 02 (2017) 299-305. Pp. 299-305.
- [6] Jamner R. Lawendatu, John S. Kekenusa, Djoni Hatidja, 2014. Regresi Linier Berganda Untuk Menganalisis Pendapatan Petani Pala. JdC, Vol. 3, No. 1. Pp.66-72
- [7] Khaifiizh Hastuti, 2012. Analisis Komparasi Algoritma Klasifikasi Data Mining Untuk Prediksi Mahasiswa Non Aktif. Seminar Nasional Teknologi Informasi & Komunikasi Terapan 2012 (Semantik 2012). Pp.241-249
- [8] Mujib Ridwan, 2013. Penerapan Data Mining Untuk Evaluasi Kinerja Akademik Mahasiswa Menggunakan Algoritma Naive Bayes Classifier. Jurnal EECCIS Vol.7, No. 1. Pp. 59-64.
- [9] Nike Syafitri , Sutardi dan L.M Tajidun, 2018. Aplikasi Forecasting Mengenai Angka Kelahiran Di Kota Kendari Menggunakan Metode Regresi Linear Berganda (Studi Kasus : Dinas Kesehatan Kota Kendari). semanTIK, Vol.4, No.1. Pp.31-38.

- [10] Slamet Pujiono , Armadyah Amborowati dan M. Suyanto, 2013. Analisis Kepuasan Publik Menggunakan Weka Dalam Mewujudkan Good Governance Di Kota Yogyakarta. JURNAL DASI ISSN: 1411-3201 Vol. 14 No. 2. Pp-45-55.
- [11] Uut Krismianto, 2015. SPSS (Statistical Package for the Social Sciens). Makalah Olah Data SPSS.Pp.1-13.
- [12] Yuli Mardi, 2013. Data Mining : Klasifikasi Menggunakan Algoritma C4.5. Jurnal Edik Informatika V2.i2. Pp.213-219.