

IMPLEMENTASI K-NEAREST NEIGHBOR UNTUK MENENTUKAN PREDIKSI PENJUALAN (STUDI KASUS : PT MAKSIPLUS UTAMA INDONESIA)

Rini Nuraini Sukmana¹, Abdurrahman², Yohanes Wicaksono³

^{1,2,3}STMIK BANDUNG

rnurainisukmana@gmail.com, yohaneswicaksono33@gmail.com

Abstrak

PT. Maksiplus Utama Indonesia adalah perusahaan berbasis mlm yang memproduksi dan menjual pupuk. Maksiplus memiliki banyak produk yang tersebar di Indonesia. Dari beberapa produk tersebut terkadang ada beberapa yang hasil penjualannya tidak sesuai harapan, dan menyebabkan kurangnya keuntungan perusahaan. Manajemen terkadang merasa kesulitan dalam menentukan apakah penjualan produk akan naik atau tidak. Untuk membantu dalam memprediksi penjualan diperlukan sebuah metode yang mampu memberikan solusi alternatif. Metode yang digunakan dalam memprediksi penjualan adalah *K-Nearest Neighbor*. Keluaran sistem disajikan dalam bentuk hasil angka perhitungan dari metode *K-Nearest Neighbor*. Hasil yang ingin dicapai dalam penelitian ini adalah perusahaan dapat terbantu dalam mengambil keputusan. Dari keluaran dari sistem yang berbentuk angka diharapkan dapat mudah dipahami dan digunakan sebagai bahan pertimbangan oleh manajemen dalam menentukan apakah suatu produk penjualannya akan naik atau tidak.

Kata kunci : prediksi penjualan, *K-Nearest Neighbor*, Maksiplus

Abstract

PT. Maksiplus Utama Indonesia is an MLM based company that produces and sells fertilizers. Maksiplus has many products spread in Indonesia. From some of these products sometimes there are some whose sales results are not as expected, and cause a lack of company profits. Management sometimes feels difficulty in determining whether product sales will rise or not. To help predict sales, a method that is able to provide alternative solutions is needed. The method used in this sales prediction is K-Nearest Neighbor. The system output is presented in the form of the calculation results from the K-Nearest Neighbor method. The results to be achieved in this study are companies can be helped in making decisions. From the output of the system in the form of numbers, it is expected to be easily understood and used as material for consideration by management in determining whether a product's sales will rise or not.

Keywords: sales predictions, *K-Nearest Neighbor*, Maksiplus.

1. Pendahuluan

1.1. Latar Belakang

Prediksi atau peramalan penjualan adalah suatu perhitungan untuk memprediksi keadaan di masa mendatang melalui pengujian keadaan di masa lalu. Prediksi hampir sama dengan klasifikasi dan estimasi, kecuali bahwa dalam prediksi nilai dari hasil akan ada di masa mendatang [1]. Meramalkan penjualan di masa mendatang artinya memperkirakan

besarnya volume penjualan, bahkan menentukan potensi penjualan dan luas pasar yang dikuasai di masa mendatang [1]. Salah satu kegunaan prediksi adalah membantu pemilik perusahaan dalam pengambilan keputusan adalah untuk menentukan fokus ke arah mana penjualan akan dilakukan sehingga diharapkan resiko kesalahan yang disebabkan

oleh kesalahan perencanaan dapat ditekan seminimal mungkin.

Metode-metode yang digunakan untuk prediksi diantaranya metode *K-Nearest Neighbor* (KNN). Metode ini merupakan metode untuk melakukan klasifikasi terhadap objek berdasarkan data pembelajaran yang jaraknya terdekat dengan objek tersebut. Keunggulan metode ini dalam memprediksi yaitu tangguh dan efektif terhadap training data yang noisinya besar

PT Maksiplus Utama Indonesia adalah perusahaan berbasis MLM yang memproduksi dan menjual pupuk organik dan suplemen propolis. Dengan berkembangnya perusahaan tentu saja diikuti dengan berkembangnya jumlah dan variasi data penjualan sehingga menimbulkan kesulitan tersendiri bagi perusahaan dalam memprediksi produk mana di yang penjualannya akan meningkat dan produk dan daerah mana yang perlu ditingkatkan penjualannya agar perusahaan semakin terus berkembang.

Berdasarkan uraian latar belakang maka diusulkanlah penelitian dengan judul **“Implementasi K-Nearest Neighbor Untuk Menentukan Prediksi Penjualan (Studi Kasus : PT Maksiplus Utama Indonesia)”**.

1.2. Rumusan Masalah

Berdasarkan latar belakang di atas maka penulis merumuskan permasalahan yaitu:

- Bagaimana sebaran penjualan produk PT Maksiplus Utama Indonesia?
- Bagaimana melakukan prediksi penjualan produk pada PT Maksiplus Utama Indonesia?
- Produk apa saja yang diperkirakan mengalami kenaikan atau penurunan pada periode berikutnya?

1.3. Tujuan Penelitian

Adapun tujuan yang ingin dicapai dalam penelitian ini adalah :

- Mengetahui pola penjualan produk di PT Maksiplus Utama Indonesia.
- Menerapkan metode algoritma K-Nearest Neighbor dalam memprediksi penjualan produk ke depannya.

- Mengetahui perkiraan kenaikan atau penurunan produk untuk diambil langkah berikutnya apakah perlu menambahkan atau mengurangi produksi atau meningkatkan promosi produk.

1.4. Batasan Masalah

Beberapa batasan masalah dalam penelitian ini adalah sebagai berikut:

- Membahas tentang prediksi penjualan produk pupuk organik dan suplemen pada PT Maksiplus Utama Indonesia.
- Proses yang diolah adalah data penjualan yang diambil langsung dari data warehouse mulai dari tahun 2016 bulan 5 hingga tahun 2019 bulan 10 menggunakan variable Nama barang, Bulan, dan Kuantitas tanpa mengalami praproses.
- Nilai K yang digunakan adalah 3.
- Data Training adalah data penjualan tahun 2017 bulan 1 hingga tahun 2019 bulan 5 dan Data Testing menggunakan data tahun 2019 bulan 6 hingga 10.
- Error margin hasil prediksi adalah 10%

1.5. Peramalan / Prediksi

Peramalan adalah ilmu memperkirakan kejadian di masa depan [4], dengan melibatkan pengambilan data masa lalu dan menempatkannya ke masa depan dengan suatu bentuk model matematis. Hakikatnya peramalan hanya merupakan perkiraan, tetapi dengan teknik-teknik tertentu, maka peramalan menjadi lebih dari sekedar perkiraan [6]. Bertujuan untuk meredam ketidakpastian sehingga didapat suatu perkiraan mendekati sebenarnya. Peramalan diklasifikasikan berdasarkan horizon waktu yang dilingkupinya yang terbagi dalam beberapa kategori [5]:

- Peramalan Jangka Pendek, meliputi jangka waktu harian atau mingguan, untuk merencanakan pembelian, penjadwalan kerja, jumlah tenaga kerja, penugasan kerja, dan tingkat produksi.
- Peramalan Jangka Menengah, mencakup hitungan bulanan atau kuartal. Bermanfaat untuk merencanakan

penjualan, perencanaan dan anggaran produksi, anggaran kas, serta menganalisis bermacam-macam rencana operasi.

- c. Peramalan Jangka Panjang, bersifat tahunan untuk merencanakan produk baru, pembelanjaan modal, lokasi atau pengembangan fasilitas, serta penelitian dan pengembangan.

Pendekatan umum teknik peramalan yaitu:

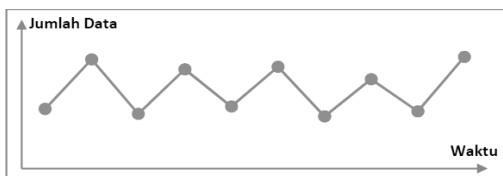
- a. Peramalan Kualitatif, menggabungkan faktor seperti intuisi, emosi, pengalaman pribadi, dan bersifat subjektif, melalui kuesioner, survey, dan riset pasar.
- b. Peramalan Kuantitatif, didasarkan pada model matematis yang beragam dengan data masa lalu (*time series*) dan variable sebab-akibat (*causal method*).

1.6. Peramalan Time Series

Merupakan Metode yang digunakan untuk menganalisis data yang merupakan fungsi dari waktu [6]. Peramalan dengan teknik ini berarti nilai masa depan diperkirakan hanya dari nilai masa lalu dan mengabaikan variabel yang lain [4].

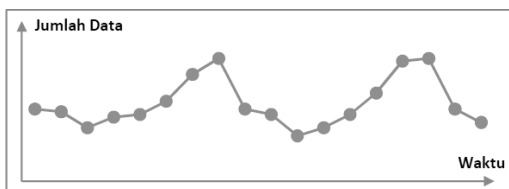
Berikut beberapa pola yang ada dalam data time series [3], yaitu:

- a. Pola Data Horizontal, terjadi bila nilai data berfluktuasi di sekitar nilai rata-rata yang konstan.



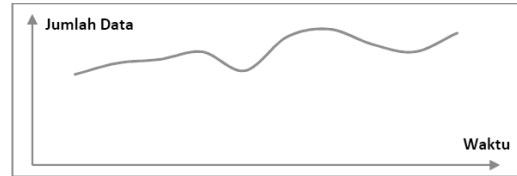
Gambar 1. Pola Data Horizontal

- b. Pola Data Musiman, terjadi jika suatu deret dipengaruhi oleh musiman.



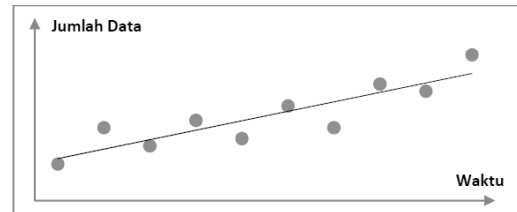
Gambar 2. Pola Data Musiman

- c. Pola Data Siklus, terjadi bila datanya dipengaruhi oleh fluktuasi ekonomi jangka panjang seperti yang berhubungan siklus bisnis.



Gambar 3. Pola Data Siklus

- d. Pola Data Tren, terjadi apabila terdapat kenaikan atau penurunan sekuler panjang dalam data.

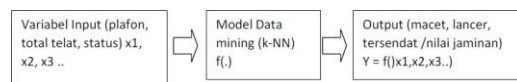


Gambar 4. Pola Data Tren

1.7. Klasifikasi

Klasifikasi adalah proses penemuan model (atau fungsi) yang menggambarkan dan membedakan kelas data atau konsep yang bertujuan agar bisa digunakan untuk memprediksi kelas dari objek yang label kelasnya tidak diketahui [2].

Klasifikasi data terdiri dari 2 langkah proses. Pertama adalah *learning*, algoritma klasifikasi dibuat untuk menganalisa data *training* lalu dipresentasikan dalam *rule* klasifikasi, dimana data tes digunakan untuk memperkirakan akurasi dari *rule* klasifikasi [2]. Klasifikasi adalah proses menempatkan objek ke dalam suatu kategori/kelas yang sudah didefinisikan sebelumnya berdasarkan model tertentu. [7]. Pemodelan Prediktif diawali dengan pembentukan pemodelan untuk memprediksi hasil.



Gambar 5. Pemodelan Prediktif

Pemodelan Deskriptif atau *clustering*, merupakan proses pengamatan terhadap kelompok data kemudian diikuti dengan

pengelompokan data/cluster terhadap data yang mempunyai kesamaan ciri.

1.8. K-Nearest Neighbor

Algoritma *K-Nearest Neighbor* (KNN) adalah sebuah metode untuk melakukan klasifikasi terhadap objek berdasarkan data pembelajaran yang jaraknya paling dekat dengan objek tersebut [10]. KN dilakukan dengan mencari kelompok objek dalam data training yang paling dekat (mirip) dengan objek pada data baru atau data testing. [8]. Algoritma KNN bersifat sederhana, bekerja dengan berdasarkan pada jarak terpendek dari sampel uji (*testing sample*) ke sampel latih (*training sample*) untuk menentukan KNN nya. Setelah mengumpulkan KNN, kemudian diambil mayoritas dari KNN untuk dijadikan prediksi dari sample uji.

KNN memiliki beberapa kelebihan yaitu tangguh terhadap *training data* yang *noise* dan efektif apabila data latih nya besar. Pada fase *training*, algoritma ini hanya melakukan penyimpanan vektor-vektor fitur dan klasifikasi data *training sample*. Pada fase klasifikasi, fitur-fitur yang sama dihitung untuk testing data atau yang klasifikasinya tidak diketahui. Jarak dari vektor baru yang ini terhadap seluruh vektor *training sample* dihitung dan sejumlah k buah yang paling dekat diambil. Titik yang baru klasifikasinya diprediksikan termasuk pada klasifikasi terbanyak dari titik-titik tersebut.

Ada banyak cara untuk mengukur jarak kedekatan antara data baru dengan data lama (data *training*), yang sering digunakan adalah *euclidean distance*, yaitu :

- D = jarak *Euclidean*
- a = sampel data
- b = data uji
- i = nilai k atau konstanta jumlah tetangga terdekat
- n = nilai atribut dimensi data

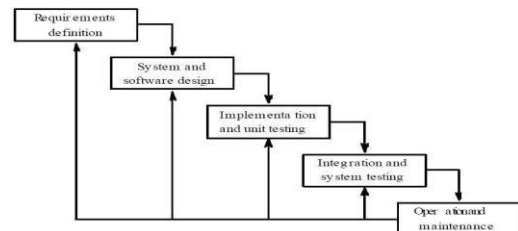
Adapun langkah-langkah menghitung algoritma KNN antara lain :

- a. Menentukan Parameter K (jumlah tetangga terdekat).
- b. Menghitung kuadrat jarak *euclidean* (*query instance*) masing –masing objek terhadap data sampel yang diberikan.
- c. Mengurutkan jarak tersebut dalam kelompok yang mempunyai Euclid terkecil
- d. Mengumpulkan kategori Y berdasarkan K atau data tetangga terdekat.
- e. Dengan kategori *Nearest Neighbor* paling mayoritas maka menghasilkan prediksi. Nilai k terbaik pada algoritma ini tergantung data. Nilai k yang tinggi akan mengurangi efek *noise* pada klasifikasi, tapi membuat batasan setiap klasifikasi semakin kabur. Nilai k yang bagus dapat dipilih dengan optimasi parameter.

2. Pembahasan

2.1. Metodologi

Pengembangan aplikasi menggunakan paradigma *waterfall*. Langkah umum metode *waterfall* [9] adalah sebagai berikut:



Gambar 6. Metode Pengembangan Waterfall

Pendekatan Waterfall melewati 5 proses, proses-proses tersebut dapat dijelaskan sebagai berikut:

- a. *Requirements Definition*, Merupakan analisis kebutuhan dan Analisis Perancangan dari sistem yang akan dibuat. Juga melakukan pengumpulan data dan penetapan kebutuhan semua elemen sistem.
- b. *System and software design*, Merupakan tahapan perancangan desain *interface* sesuai kebutuhan dari permasalahan yang telah dikemukakan sebelumnya dengan meninjau bagian perumusan masalah dan latar belakang.

- c. *Implementation and Testing Unit*, Di bagian ini aplikasi akan diterapkan dan diuji oleh pengguna atau tester.
- d. *Integration and testing system*, Pada tahapan ini sistem akan diintegrasikan kepengguna sistem / user.
- e. *Operational maintenance*, Pada tahap akhir yaitu aplikasi dioperasikan pada perusahaan serta melakukan perawatan / *maintenance* secara berkala.

2.2. Pembahasan

Pada bagian ini akan dibahas mengenai hasil implementasi dan pembahasan hasil prediksi menggunakan algoritma KNN. dicontohkan dengan kasus memprediksi penjualan produk Maxigrow bulan ke 11 dengan menggunakan data bulan sebelumnya dengan nilai $K=3$.

- a. Data Selection, dengan sampel data diambil dari data penjualan produk MaxiGrow dengan kasus penjualan bulan sebelumnya digunakan untuk meramalkan penjualan bulan ke 11, sampel data seperti tabel di bawah

Tabel 1. Sampel Data

Bulan	Jumlah
1	470
2	424
3	388
4	488
5	401
6	452
7	444
8	398
9	439
10	411
11	???

- b. Transformation, Data tersebut kemudian akan digunakan untuk training yang dikelompokkan menjadi 2 bagian yaitu data *input* dan data target. Data *input* merupakan data penjualan dari bulan ke-1 sampai bulan ke-5, sedangkan data target menggunakan data bulan ke-6, dilanjutkan data *input* bulan ke-2 sampai bulan ke-6 dan data target menggunakan data bulan ke-7 dan seterusnya. Dan akhirnya data bulan ke 7 hingga bulan

ke 11 menjadi Data Testing. Hasil dapat dilihat pada tabel di bawah:

Tabel 2. Hasil Data Training

No	Data Input					Target
1	470	424	388	488	401	452
2	424	388	488	401	452	444
3	388	488	401	452	444	398
4	488	401	452	444	398	439

Tabel 3. Hasil Data Testing

452	444	398	439	411	?
-----	-----	-----	-----	-----	---

- c. Menghitung jarak *Euclidean*

Tabel 4. Hasil Perhitungan Jarak *Euclidean*

No	Perhitungan <i>euclidian distance</i>	Hasil
1	$\sqrt{(470-452)^2 + (424-444)^2 + (388-398)^2 + (488-439)^2 + (401-411)^2}$	57.6628
2	$\sqrt{(424-452)^2 + (388-444)^2 + (488-398)^2 + (401-439)^2 + (452-411)^2}$	123.0650
3	$\sqrt{(388-452)^2 + (488-444)^2 + (401-398)^2 + (452-439)^2 + (444-411)^2}$	85.4342
4	$\sqrt{(488-452)^2 + (401-444)^2 + (452-398)^2 + (444-439)^2 + (398-411)^2}$	79.0886

- d. Pengurutan Jarak Terkecil

No	Hasil	Urutan	Termasuk nearest neighbor (K)
1	57.6628	1	Ya
2	123.0650	4	Tidak
3	85.4342	3	Ya
4	79.0886	2	Ya

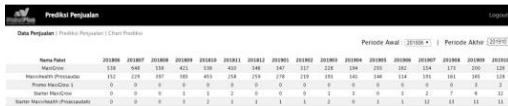
- e. Menghitung rata-rata nilai *Object* pada jangkauan K, Dari hasil di atas disimpulkan bahwa data yang digunakan adalah nomor 1 (tetangga = 452), 3 (tetangga = 398) dan 4 (tetangga = 439), sehingga dapat dihitung nilai prediksi untuk bulan 11 adalah:

$$(452+439+398)/3 = 1251/3 = \mathbf{439}$$

Berikut adalah tampilan dari aplikasi yang telah dibuat



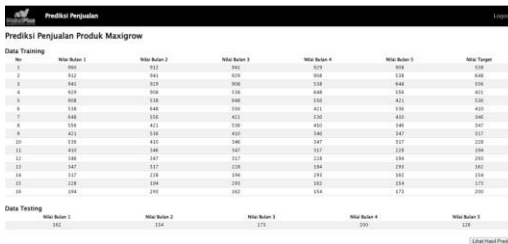
Gambar 7. Halaman Login



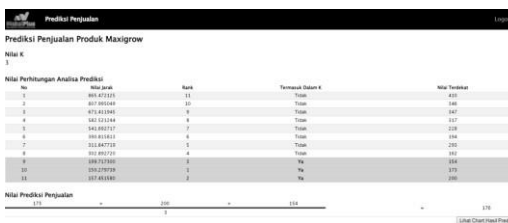
Gambar 8. Halaman Laporan Penjualan



Gambar 9. Halaman Prediksi Pejualan Awal



Gambar 10. Halaman Prediksi Data Training dan Data Testing

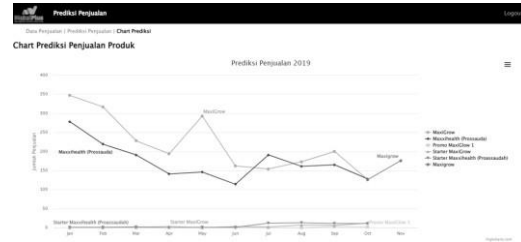


Gambar 11. Halaman Hasil Prediksi

3. Kesimpulan

Setelah dilakukan analisis, perancangan dan pembuatan aplikasi, maka dapat diambil kesimpulan sebagai berikut:

- a. Aplikasi yang dibuat sudah dapat menampilkan tampilan hasil dan pola penjualan tiap produk yang dipilih.
- b. Aplikasi yang dibuat sudah dapat



Gambar 11. Halaman Chart Hasil Prediksi

Setelah aplikasi siap digunakan maka dilakukan evaluasi untuk mengukur tingkat presisi dari hasil prediksi. Evaluasi tersebut dapat dilihat pada tabel dibawah ini.

Produk	Periode	Nilai Real	Hasil Prediksi	Selisih (%)
MaxiGrow	201901	470	456	2.98%
MaxiGrow	201902	424	431	1.65%
MaxiGrow	201903	388	409	5.41%
Maxigrow	201904	488	413	15.37%
MaxiGrow	201905	401	400	0.25%
MaxiGrow	201906	452	449	0.66%
MaxiGrow	201907	444	438	1.35%
MaxiGrow	201908	398	426	7.04%
MaxiGrow	201909	439	438	0.23%
MaxiGrow	201910	411	412	0.24%
Maxxihealth	201901	380	369	2.89%
Maxxihealth	201902	328	323	1.52%
Maxxihealth	201903	373	364	2.41%
Maxxihealth	201904	331	323	2.42%
Maxxihealth	201905	322	332	3.11%
Maxxihealth	201906	361	361	0.00%
Maxxihealth	201907	341	335	1.76%
Maxxihealth	201908	356	351	1.40%
Maxxihealth	201909	337	344	2.08%
Maxxihealth	201910	333	338	1.50%

Berdasarkan hasil prediksi data di atas maka dapat diambil kesimpulan bahwa *error margin* hasil dari prediksi ini dapat dikatakan di bawah 10%. Tingkat akurasi ini dapat diperbaiki dengan menambah jumlah data *training* dan melakukan praprocess data, sehingga kasus baru yang akan diprediksi penjualannya sudah ada nilai acuannya di dalam *database*.

- c. menampilkan tampilan hasil prediksi penjualan tiap produk yang dipilih dengan menerapkan algoritma *K- Nearest Neighbor*.
- d. Tampilan aplikasi telah dapat menampilkan produk mana saja yang diperkirakan naik atau turun pada periode berikutnya berdasarkan tampilan *line chart* sehingga dapat diambil keputusan berikutnya yang apabila produk diprediksi mengalami penurunan maka perlu diadakan promosi untuk meningkatkan

penjualan dan apabila ada prediksi kenaikan maka perlu ditambah untuk produksi produk di pabrik.

4. Saran

Sesuai dengan hasil analisis dan evaluasi terhadap sistem, saran yang dapat diberikan untuk pengembangan lebih lanjut adalah sebagai berikut:

- a. Aplikasi dapat dikembangkan agar aplikasi dapat dihubungkan dengan aplikasi website utama.
- b. Sistem dapat dikembangkan lagi untuk memberikan data hasil prediksi secara real-time dengan data terbaru.
- c. Data awal yang diperoleh dari *data warehouse* diolah terlebih dahulu atau dilakukan praproses sehingga bisa memperkecil error margin hasil prediksi.

Erlangga.

- [9] WIDIARSANA, O., PUTRA, N.W., BUDIYASA, P.G.I., BISMANTARA, A.N.I., MAHAJAYA, S.N. (2011) *Data Mining: Metode Clasification K-Nearest Neighbor (KNN)*. Bali: Program Studi Teknologi Informasi Universitas Udayana.

DAFTAR PUSTAKA

- [1] R. N. Sukmana, R. K. Agustini and M. FransiscaSiregar, "APLIKASI INVENTARIS DENGAN PENGIMPLEMENTASIAN ALGORITMA APRIORI UNTUK REKOMENDASI PENGADAAN BARANG (STUDI KASUS: SMK BELA NUSANTARA)," *Jurnal Teknologi Informasi dan Komunikasi*, pp. 50-59, 2017.
- [2] ERIYANTO OCKI (2012). Analisis Peramalan Penjualan *Handphone* Blackberry Pada PT. Selular Shop Mall
- [3] HAN, JIAWEI DAN KAMBER, MICHELINE. (2006), *Data Mining : Concept and Techniques Second Edition*. Morgan Kaufmann Publishers.
- [4] HARTANTO, R. D. (2012) Pengaruh Komunikasi, Kepemimpinan dan Budaya Organisasi terhadap Kinerja karyawan Pada PDAM Kabupaten Demak. Semarang: Fakultas Ekonomi Universitas Stikubank.
- [5] ISHAK, AULIA. (2010) Manajemen Operasi. Yogyakarta: Graha Ilmu.
- [6] PANDIE, EMERENSYE S. Y. (2012) Implementasi Algoritma Data mining K-Nearest Neighbour (KNN) Dalam Pengambilan Keputusan Pengajuan Kredit. Jurusan Ilmu Komputer, Fakultas Sains dan Teknik, Kupang: Universitas Nusa Cendana.
- [7] SAXENA, KRATI. (2014) *Diagnosis of Diabetes Mellitus using K Nearest Neighbor Algorithm. Internasional Journal of Computer Science Trends and Technology (IJCSST) ISSN : 2347-8578 Volume 2 Issue 4, July-Agus 2014.*
- [8] SOMMERVILLE, IAN. (2003) *Software Engineering*. Edisi 6 Jilid 1, Jakarta:

