

## **Penerapan Data Mining Untuk Prediksi Layanan Produk Indihome Menggunakan Metode K-Nearest Neighbor**

*The Data Mining Application for IndiHome Product Service Prediction by Using K-Nearest Neighbor Method*

Arwa Ulayya Haspriyanti<sup>1</sup>, Putri Taqwa Prasetyaningrum<sup>2</sup>  
<sup>1,2</sup>Program Studi Sistem Informasi, Fakultas Teknologi Informasi,  
Universitas Mercu Buana Yogyakarta, Indonesia  
Email: alyalfa96@gmail.com<sup>1</sup>, putri@mercubuana-yogya.ac.id<sup>2</sup>

### **ABSTRAK**

PT. Telkom adalah perusahaan yang bergerak dalam bidang telekomunikasi dengan produk IndiHome yang merupakan gabungan dari jenis produk dan layanan telecommunication, information, media dan edutainment menjadi satu terdiri dari layanan internet, internet voice dan internet interaktif (UseTV cable) atau Triple Play. Pengembangan strategi pemasaran yang maksimal memerlukan bauran pemasaran efektif dan efisien. Pada prinsipnya, strategi pemasaran memberikan arah dalam kaitannya dengan variabel-variabel seperti segmentasi pasar, target, positioning, dan elemen bauran pemasaran. Untuk menunjang penelitian ini, maka variabel yang digunakan adalah variabel operasional untuk analisis eksternal dan internal. Dalam penelitian ini metode yang digunakan adalah metode penelitian kualitatif. Sedangkan jenis penelitiannya adalah deskriptif. Data primer diperoleh langsung dari subjek penelitian melalui observasi, wawancara, dan dokumentasi. Dilihat dari banyaknya permintaan konsumen akan produk layanan berdasarkan data penjualan selama 1 tahun terakhir maka dibutuhkan prediksi untuk penjualan produk layanan terlaris, guna untuk mempermudah pihak perusahaan dalam perencanaan penyediaan layanan. Untuk mengetahui penjualan produk layanan terlaris digunakan teknik klasifikasi data mining dan algoritma *K-Nearest Neighbor*. Hasil dari penelitian ini adalah prediksi penjualan layanan terlaris sebanyak 1 jenis produk dari 3 jenis produk yang terjual yaitu Internet, Internet Voice, Internet Voice Usestv. Berdasarkan nilai akurasi terhadap klasifikasi penjualan produk terlaris sebesar 99,99 %.

**Kata kunci:** telekomunikasi, internet, *K-Nearest Neighbor*

### **ABSTRACT**

PT. Telkom is a company engaged in the field of telecommunication that develops IndiHome product. IndiHome itself is the combination of product and service types including telecommunication, information, media and edutainment becomes one consisting internet service, internet voice and interactive internet (UseTV cable) or Triple Play. The development of a maximum marketing strategy requires an effective and efficient marketing mix. In principle, the marketing strategy provides direction in relation with variables such as market segmentation, target, positioning, and marketing mix element. In order to support this study, the variable used is operational variable for external and internal analysis. In this study, the method used is qualitative research method. Meanwhile, the type of the study is descriptive. The primary data is directly obtained from the research subject through observation, interview, and documentation. As seen from the large number of customer request related to the service product based on sales data for the past 1 year, it is important to get the sales prediction of best-selling service product, in order to facilitate the company in developing service provision planning. Furthermore, the use of data mining classification technique and K-Nearest Neighbor algorithm is needed to find out the best-selling service product sales. The result of this study is the prediction of the best-selling service product sales which is 1 from the total of 3 product types that are sold including Internet, Internet Voice, and Internet Voice Usestv. It is based on the accuracy value towards the classification of the best-selling product sales of 99.99%.

**Keywords:** telecommunications, internet, K-Nearest Neighbor.

## 1. PENDAHULUAN

### Penerapan Data Mining Untuk Prediksi Layanan Produk Indihome Menggunakan Metode K-Nearest Neighbor

Telkom merupakan perusahaan yang menyediakan jasa dalam bidang telekomunikasi terkemuka di Indonesia, dengan demikian Telkom wajib memberikan layanan terbaik kepada pelanggannya. Laju perkembangan teknologi akses jaringan dan teknologi berbasis internet memungkinkan bagi operator layanan untuk dapat menyediakan layanan yang beragam (*multi service*) seperti layanan suara (telepon), data (internet) dan gambar (*IP-TV*) dalam satu jaringan akses dengan layanan *triple-play*. IndiHome merupakan layanan *Triple Play* dari Telkom yang terdiri dari *Internet on Fiber* atau *High Speed Internet*, *Phone (Telepon Rumah)*, dan *IPTV (UseTV Cable)*. (Telkom, 2019)

Salah satu permasalahan yang dihadapi oleh perusahaan yang bergerak dibidang penjualan jasa adalah mempertahankan layanan-layanan tersebut. Dengan melakukan prediksi dapat membantu pihak Telkom untuk mengetahui apakah pelanggan Telkom nyaman terhadap layanan produk yang disediakan yaitu Indihome. Layanan dapat dikatakan baik ketika banyak pelanggan yang berminat dan membutuhkan layanan Indihome sesuai dengan kebutuhan masing-masing.

Prediksi atau peramalan penjualan (*forecasting*) adalah suatu perhitungan untuk meramalkan keadaan di masa mendatang melalui pengujian keadaan di masa lalu. Meramalkan penjualan di masa mendatang berarti menentukan perkiraan besarnya volume penjualan, bahkan menentukan potensi penjualan dan luas pasar yang dikuasai di masa yang akan datang (Eriyanto, 2012). Salah satu dari kegunaan prediksi adalah untuk membantu pemilik perusahaan dalam mengambil keputusan dalam menentukan jumlah barang yang harus disediakan oleh perusahaan. Selain itu prediksi dapat membantu pihak perusahaan dalam perencanaan penyediaan layanan, karena prediksi ini dapat memberikan *output* terbaik sehingga diharapkan resiko kesalahan yang disebabkan oleh kesalahan perencanaan dapat

ditekan seminimal mungkin. Prediksi biasanya digunakan untuk menemukan informasi dari sejumlah data yang besar sehingga diperlukan *data mining*.

*K-Nearest Neighbor* adalah suatu metode yang menggunakan algoritma supervised dimana hasil dari query instance yang baru diklasifikasikan berdasarkan mayoritas dari label class pada KNN. Tujuan dari algoritma KNN adalah mengklasifikasikan objek baru berdasarkan atribut dan training data. KNN juga merupakan metode yang baik digunakan untuk melakukan suatu prediksi.

Oleh karena itu, berdasarkan permasalahan diatas, peneliti tertarik untuk melakukan penelitian dengan judul "Penerapan Data Mining Untuk Prediksi Penjualan Produk Indihome Terlaris Menggunakan Metode K-Nearest Neighbor" dengan cara melakukan pengolahan data dan penerapan data mining dengan memanfaatkan data-data terdahulu sehingga dapat menghasilkan informasi dan rangkaian pola yang bisa digunakan untuk memprediksi variabel lain yang belum diketahui nilai ataupun jenisnya guna memecahkan permasalahan tersebut.

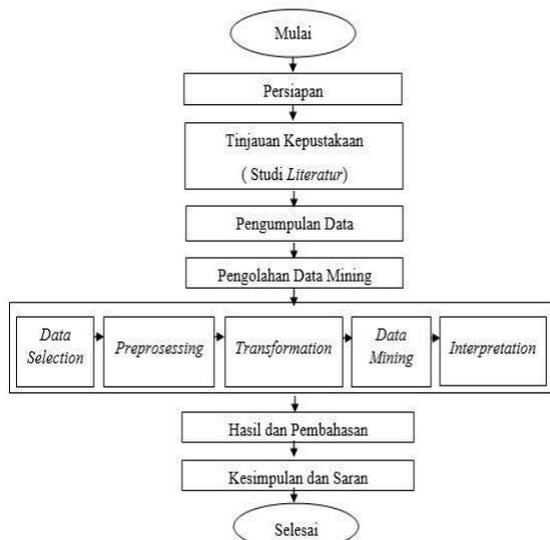
## 2. TINJAUAN PUSTAKA

Selanjutnya adalah penelitian yang dibuat yaitu untuk memperkirakan produk layanan Indihome terlaris dengan menggunakan metode K-Nearest Neighbor. Perbedaan dengan penelitian sebelumnya adalah lokasi penelitian yang akan dikerjakan dan atribut yang digunakan yaitu wilayah pemasangan, package name, jenis layanan, serta rumusan masalah yang berbeda, untuk kasus saat ini di kerjakan dengan tujuan untuk mengetahui produk layanan indihome terlaris. Alasan pemilihan atribut yaitu efisiensi dari data dan hasil akan mempermudah manager dalam mengetahui layanan mana yang dapat di pertahankan untuk produksi tahun berikutnya. Sedangkan metode K-Nearest Neighbor atau k-NN adalah algoritma yang berfungsi untuk melakukan klasifikasi suatu data berdasarkan data pembelajaran (*train data sets*),

yang diambil dari  $k$  tetangga terdekatnya (*nearest neighbors*). Dengan  $k$  merupakan banyaknya tetangga terdekat.

### 3. METODOLOGI PENELITIAN

Tahapan penelitian yang dituangkan dalam diagram alir ini. Menggambarkan proses penelitian yang akan ditempuh sekaligus menggambarkan penelitian secara keseluruhan. Tahapan yang akan ditempuh yaitu:



**Gambar 1. Tahapan Penelitian**

Selanjutnya Penentuan nilai  $k$ . Penentuan nilai  $k$  yang digunakan dalam klasifikasi tidak memiliki aturan yang baku, namun pada penelitian ini nilai  $k$  yang digunakan adalah 3, 5, 7, 9, 11, 13, 15.

Perhitungan jarak antar data *training* dan data uji (*test*). Teknik perhitungan jarak yang digunakan dalam metode KNN ini adalah Jarak *Euclidean Distance*. Dengan rumus sebagai berikut

$$\sqrt{\sum_{i=1}^K (X_i - Y_i)^2}$$

Pengurutan data hasil perhitungan. Jarak yang telah didapatkan kemudian diurutkan dari yang paling dekat jaraknya sampai yang paling jauh (*ascending*).

Menentukan kelompok data hasil uji berdasarkan label mayoritas dari  $k$  tetangga terdekat.

### 4. PEMBAHASAN

Setelah melakukan analisa terhadap perancangan dengan tahapan *data mining* untuk menghasilkan prediksi penjualan produk layanan terlaris pada PT.Telkom Indonesia di Yogyakarta dengan menggunakan Algoritma *K-Nearest Neighbor*, analisa ini berakhir dengan melakukan proses *data mining* yang sesungguhnya, maka hasil yang dicapai oleh peneliti adalah untuk mengetahui prediksi penjualan layanan terlaris berdasarkan data penjualan produk layanan dari data penjualan di tahun 2019 yang ada di PT.Telkom Indonesia di Yogyakarta. Peneliti menggunakan aplikasi *RapidMiner* untuk mempermudah proses *data mining*, yang menghasilkan informasi prediksi penjualan produk indihome terlaris. Dan didapatkan hasil prediksi penjualan produk terlaris sebanyak 3 jenis produk diantaranya Layanan 1P (Internet), Layanan 2P (Internet Voice), Layanan 3P (Internet Voice Usestv).

Data yang digunakan dalam penelitian ini yaitu data penjualan produk layanan berdasarkan penjualan 1 tahun terakhir yaitu tahun 2019 yang berasal dari PT.Telkom Indonesia regional Yogyakarta, dengan data sebagai berikut :

**Table 1. Data Selection**

No	Nama Produk	Kuantitas	Last Update
1	INT ERNET	1	January 2019
2	INT ERNET	1	January 2019
3	INT ERNET	1	January 2019
4	INT ERNET	1	January 2019
5	INT ERNET	1	January 2019
6	INT ERNET	1	January 2019
7	INT ERNET	1	January 2019
8	INT ERNET	1	January 2019
9	INT ERNET	1	January 2019
10	INT ERNET	1	January 2019

**Table 1** Merupakan data yang telah diselection dan siap digunakan untuk proses data mining. Jumlah data yang di proses untuk pengujian adalah 39.631 data.

**Table 2. Preprocessing 1**

Jenis Layanan	Januari	Februari	Maret	April
Single Play	0	8	1	4
Single Play	5	4	0	2
Single Play	6	2	0	4
Single Play	5	1	0	5
Single Play	1	1	4	3
Single Play	0	0	3	3
Single Play	5	5	4	2
Single Play	5	0	1	4
Single Play	4	6	2	3

**Tabel 2** Merupakan hasil pengelompokan penjualan produk layanan single play untuk setiap bulan dengan total penjualan sebanyak 967 item.

**Table 3. Preprocessing 2**

Jenis Layanan	Januari	Februari	Maret	April
Dual Play	12	48	56	57
Dual Play	40	53	56	76
Dual Play	51	33	36	46
Dual Play	58	60	68	67
Dual Play	65	41	77	53
Dual Play	36	69	71	46
Dual Play	76	63	63	28
Dual Play	73	42	61	46
Dual Play	60	52	64	52

**Tabel 3** Merupakan hasil pengelompokan penjualan produk LED untuk setiap bulan dan tahun dengan total penjualan sebanyak 19518 item.

**Table 4. Preprocessing 3**

Jenis Layanan	Januari	Februari	Maret	April
Triple Play	19	32	27	26
Triple Play	37	47	39	30
Triple Play	42	22	25	29
Triple Play	41	43	45	50
Triple Play	50	24	34	23
Triple Play	30	52	28	43
Triple Play	37	41	38	30
Triple Play	59	23	27	32

**Tabel 4** merupakan hasil pengelompokan penjualan produk Layanan Internet Voice Usee

Tv untuk setiap bulan dengan total penjualan sebanyak 19146 item.

Selanjutnya Perhitungan jarak antar data *training* dan data uji (*test*) yang ada pada tahap *transformation* dengan menggunakan perhitungan *Euclidean Distance* sebagai berikut :

$$d1 = \sqrt{(0-3)^2 + (8-2)^2 + (1-4)^2 + (4-2)^2 + (2-2)^2} = 1,7222$$

$$d2 = \sqrt{(5-3)^2 + (4-2)^2 + (0-4)^2 + (2-2)^2 + (6-2)^2} = 2.0466$$

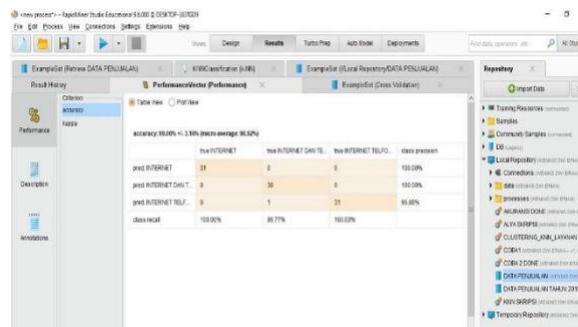
$$d3 = \sqrt{(6-3)^2 + (2-2)^2 + (0-4)^2 + (4-2)^2 + (0-2)^2} = 1,927$$

$$d4 = \sqrt{(5-3)^2 + (1-2)^2 + (0-4)^2 + (5-2)^2 + (4-2)^2} = 1,9066$$

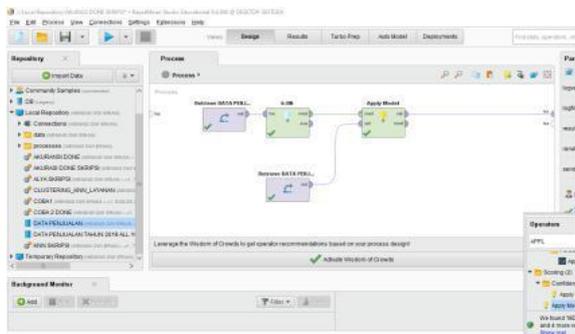
$$d5 = \sqrt{(1-3)^2 + (1-2)^2 + (4-4)^2 + (3-2)^2 + (0-2)^2} = 2.0588$$

Pengurutan data hasil perhitungan. Jarak yang telah didapatkan kemudian diurutkan dari yang paling dekat jaraknya sampai yang paling jauh (*ascending*).

Maka di peroleh hasil sebagai berikut pada implementasi rapid miner :



**Gambar 2. Akurasi Data sebesar 99,99%**



**Gambar 3. Susunan operator algoritma k-NN**

Row No.	TARGET	prediksi/TARGET	JENIS LAYANAN
2	2.912	2.912	INTERNET
3	2.939	2.939	INTERNET
4	2.976	2.976	INTERNET
5	2.772	2.772	INTERNET
6	2.248	2.248	INTERNET
7	3.959	3.959	INTERNET
8	2.729	2.729	INTERNET
9	3.790	3.790	INTERNET
10	3.173	3.173	INTERNET
11	3.900	3.900	INTERNET
12	4.095	4.095	INTERNET
13	4.008	4.008	INTERNET

**Gambar 4. Hasil prediksi layanan single play sebesar 5,028**

Row No.	TARGET	prediksi	JENIS LAYANAN
12	14	13.981	INTERNET VOICE
17	13	14.382	INTERNET VOICE
21	15	15.454	INTERNET VOICE
14	21	17.505	INTERNET VOICE
9	19	18.153	INTERNET VOICE
3	37	22.927	INTERNET VOICE
10	18	22.451	INTERNET VOICE
7	49	28.197	INTERNET VOICE
5	30	39.844	INTERNET VOICE
4	36	39.756	INTERNET VOICE
2	43	32.919	INTERNET VOICE
6	32	22.994	INTERNET VOICE
1	54	37.427	INTERNET VOICE

**Gambar 5. Hasil prediksi layanan single play sebesar 37,427**

Row No.	TARGET	prediksi	JENIS LAYANAN
21	71	70.467	INTERNET TV
19	197	80.105	INTERNET TV
12	77	80.172	INTERNET TV
20	81	83.251	INTERNET TV
17	152	83.633	INTERNET TV
18	95	83.826	INTERNET TV
26	79	85.285	INTERNET TV
28	91	86.184	INTERNET TV
27	90	88.113	INTERNET TV
25	91	90.890	INTERNET TV
24	82	90.187	INTERNET TV
31	104	90.511	INTERNET TV

**Gambar 6. Hasil prediksi layanan single play sebesar 90,511**

#### 4. KESIMPULAN

Pengubahan data primer yang telah diuji menggunakan rapid miner menghasilkan akurasi data sebesar 99,99%, dimana data tersebut dapat digunakan sebagai acuan prediksi pada proses atau tahapan selanjutnya

Implementasi data mining pada rapid miner menghasilkan 1 jenis produk layanan terlaris dari 3 produk layanan, dengan perhitungan sebagai berikut : layanan (*internet;5,028*), (*internet voice;37,427*), (*internet voice usestv;90,511*)

Pengolahan data sehingga menghasilkan prediksi paling akurat hanya dapat dilakukan dengan menggunakan metode K-Nearest Neighbor karena pengujian telah membuktikan bahwa akurasi data lebih tinggi dibandingkan dengan metode lainnya.

#### 5. UCAPAN TERIMA KASIH

Terimakasih kepada seluruh pihak yang telah membantu proses penelitian sehingga dapat terselesaikan dengan baik dan tepat waktu. Semoga penelitian ini dapat bermanfaat untu penelitian yang akan datang, sehingga dapat juga ditambahkan penerapan data mining dengan menggabungkan dua metode algoritma data mining sehingga akan menghasilkan hasil yang lebih variatif dan tentunya bernilai informasi yang sangat tinggi

#### DAFTAR PUSTAKA

Ginting, Selvia Lorena Br dkk. 2014. *Teknik Data Mining Untuk Memprediksi Masa Studi Mahasiswa Menggunakan Algoritma K-Nearest Neighbor*. Jurnal Teknik Komputer Unikom. Volume: 3, No.2.

Hermawan, Ferry dan Agung Halim. 2017. *Implementasi Metode K-Nearest Neighbor Pada Aplikasi Data Penjualan PT. Multitek Mitra Sejati*.

Hutami, Resti dan Erna. 2016. *Implementasi Metode K-Nearest Neighbor Untuk Prediksi Penjualan Furniture Pada CV. Octo Agung Jepara*.

Indrajani. 2015. *Database Design (Case Study All in One)*. Jakarta:PT.Elex Media Komputindo

Jogiyanto, HM. 2008. *Metodologi Penelitian Sistem Informasi: Pedoman dan Contoh*

*Melaksanakan Penelitian dibidang Teknologi Informasi.* Yogyakarta:Andi.

Kusrini, Lutfi Emha Taufiq. 2009. *Algoritma Data Mining.* Andi:Yogyakarta .  
Larose, Daniel T. 2005. *Discovering Jurnal Multimedia & Artificial Intelligence, Volume X, Nomor X, Februari 2019*

*Knowledge in Data : An Introduction to Data Mining.* John Willey's & Sons, Inc.

Nanja, Muis dan Purwanto. 2015. *Metode K-Nearest Neighbor Berbasis Forward Selection Untuk Prediksi Harga Komoditi Lada.* Jurnal Pseudocode.Vol.2, No.1 ISSN 2355 – 5920.

Nazir, M. 2003. *Metode Penelitian.* Jakarta:Ghalia Indonesia.

Novianti, Andi Gita & Prasetyo D. 2017. *Penerapan Algoritma K-Nearest Neighbor (KNN) Untuk Prediksi Waktu Kelulusan Mahasiswa.* Seminar Nasional APTIKOM.

Sulianta, Feri dan Dominikus Juju. 2010. *RapidMiner (YALE).* PT.Elex Media Komputindo.

Sumarlin. 2015. *Implementasi Algoritma K-Nearest Neighbor Sebagai Pendukung Keputusan Klasifikasi Penerima Beasiswa PPA dan BBM.* (<http://ejournal.undip.ac.id/index.php/jsinbi>) diakses 3 Maret 2018.

Kurniawan, 2015. *Software Aplikasi Data Mining*

Telkom, 2019. *Profil PT.Telekomunikasi Indonesia*

