



# K-NEAREST NEIGHBOR ALGORITHM AS A PREDICTION METHOD BEST SELLING ELECTRONIC PRODUCTS

Afrizal Martin<sup>1</sup>, Ari Bowo<sup>2</sup>, Eviliana Putri Dwifandi<sup>3</sup>

<sup>1</sup>Digital Bussines of Study Program, Bakti Nusantara Institute

<sup>2,3</sup>Information Systems of Study Program, Bakti Nusantara Institute

<sup>1</sup>Jl. Pramuka, Labuhan Ratu II, Way Jepara, Lampung Timur, Indonesia

<sup>2,3</sup>Jl. Wisma Rini No. 09 Pringsewu, Lampung, Indonesia

\*Corresponding author

[afrizalmartin@ibnus.ac.id](mailto:afrizalmartin@ibnus.ac.id)

[aribowo@ibnus.ac.id](mailto:aribowo@ibnus.ac.id),

[eviliana](mailto:eviliana)

[putridwifandi@gmail.com](mailto:putridwifandi@gmail.com)

## Article history:

Received October 22, 2024

Revised November 8, 2024

Accepted November 14, 2024

## Keywords:

Data Mining;

Prediction;

K-Nearest Neighbor;

Orange.

## Abstract

*The use of electric vehicles is now supported by a braking Development increasingly advanced technology rapid bring impact on various sectors, including sector trade. One of the affected sectors impact is trading product electronics. Products electronic be one of the most popular products among the public Because need will increasingly advanced technology increased. MK Store is having trouble in predict sale product electronic best seller, then from That done study optimization planning provision stock product electronics at MK Store with use prediction sale product electronic best seller. With implementing prediction models sale product electronic best seller at MK Store using data mining methods for optimize planning supplies stock product electronics at MK Store with use prediction sale product electronic best seller, and implementing a prediction model sale product electronic best seller at MK Store developed use K-Nearest Neighbor method in provision product electronics. Using The K-Nearest Neighbor algorithm is believed to can give very accurate results with count use distance neighbor nearest , and use variable 10 variables For predicted . Using the Orange tool because can make it easier in operate with use between drag-and-drop face create channel work, data analysis and own diverse widgets components that include data processing, modeling predictive, visualization and analysis text. The result of research conducted use the K-Nearest Neighbor algorithm produces mark with correlation positive (excellent), and the results prediction product electronic best seller at MK Store namely Magicom/Rice Cooker. With existence study This expected for study next, recommended expand range with develop criteria new and considering use method alternative.*

## 1.0 INTRODUCTION

Development technology and information that is increasingly rapid bring impact on various sectors, including sector trade. One of the sector affected trade impact is trading product electronics. Products electronic be one of the most popular products among the

public Because need will increasingly advanced technology increased. MK Store as one of the shop electronics in the Regency Pringsewu need do prediction sale product electronic best seller for make it easier planning supply stock goods. Based on data from the Central Statistics Agency (BPS), it was recorded Product Gross Domestic Product (GDP) at Constant Prices (ADHK) Industrial products electronics and equipment electricity amounting to IDR 25.24 trillion in the first quarter of 2023. This value increased by 12.78% annual until reached Rp49.87 trillion. Growth performance industry electronic along with increasing request abroad. It looks like from mark export IT industry increased by 9.42% [1].

In research this , refers to other research as reference , based on research conducted by [2, p . 12] . Research This review use K-Nearest Neighbor method in prediction sale frame. Done sales data analysis for prediction sale frame with use Data Mining techniques. In implementing Data Mining using K-Nearest Neighbor method for predict, obtained results by 88.89% when using sales data and 80.0% when using used data. Research furthermore done by [3] Study This utilizing Data Mining for analyze sale feed the most sought after animal . Done analysis use the initial data set consists of from 360 entries sale feed animals. After the data cleaning process, the number reduce into 208 entries. Data division is done use split method with ratio of 80% for training data and 20% for testing data. Calculation process done manually with use RapidMiner device. After set K value =5, analysis continued. Calculation done use *Euclidean Distance*, and its results sorted based on distance the smallest for do classification category product. Product Then classified into 6 products that are less best-selling and 35 best- selling products, with level accuracy reached 80.4%. Research furthermore conducted by [4, p. 10] . Research discuss about method *K-Nearest Neighbor* For predict sale best seller at Zerita Furniture Store. Through Data Mining calculations with technique classification and K-Nearest Neighbor algorithm, obtained prediction that cupboard clothes become product furniture best seller, with sale reached 90 units. Prediction sale product furniture for 2020 is done use application *RapidMiner*. Study furthermore conducted by [5]. Research This discuss about method *K-Nearest Neighbor* is used for predict sale price onion red in Yogyakarta. Research the use algorithm *K-Nearest Neighbor* For analyze four condition testing. Accuracy results the highest obtained was 91.67% and resulted in mark neighbor closest with produce level accuracy highest namely K=5. In the data modeler in the test conditions second of the 4 conditions applied. Research furthermore conducted by [6]. Research This discuss about method *K-Nearest Neighbor* is used for prediction sale food light at Indah Jaya Store. With apply K-Nearest Neighbor algorithm with K value = 3, found that algorithm the capable predict amount sale food light for 6 months to front, especially for skin buffalo, with total sales as many as 585 units.

Study previously use algorithm *K-Nearest Neighbor* and applications *Data Mining RapidMiner*. According to study previously use application *RapidMiner* can make it easier in build channel Work data analysis without need write code. Difference with research that will be done that is with using application *Data Mining Orange* and using algorithm *K-Nearest Neighbor*. Application *Data Mining Orange* use between advance *drag and drop* so that make it easier for make channel Work data analysis and own diverse component (*widgets*) which include data processing, modeling predictive, visualization and analysis text. Using algorithm *K-Nearest Neighbor* No make assumption certain about data distribution, capabilities for adapt K value (neighbor closest), making it powerful tool in data analysis, and using 10 variables for tested its validity and reliability.

MK Store faces problem in increase sale product electronics, as well as face difficulty in to provide product best - selling electronics. Of the many product electronics sold at MK Store from various type type *brand* product result in the store is having difficulties in understand type best - selling products. Researchers will try for finish problem the with predict sale product electronic best seller at MK Store make it easier in planning supply stock product best-selling electronics. Benefits in predict sale product electronic best seller can make it easier the Store in planning stock goods. And optimize supply goods so that can increase satisfaction customer with always availability stock items in the store.

## 2.0 THEORETICAL

### 2.1. Study Previous

Study previous own role important in development study next. Using study previous for inspect weaknesses and strengths from research that has been done previously. With Thus,

research previous can made into as runway relevant and supporting theories study which will conducted. In addition, research previous done Because existence reference relevant scientific so that study more accurate. Here is study previously, which was made reference, as following:

Table 1 Research Previous

No	Title Research	Research methods	Findings Study
1	Implementation of the <i>K-Nearest Neighbor</i> Method for Prediction Sale Product Best Seller at Indah Jaya Store. [6]	<i>K-Nearest Neighbor, RapidMiner, Data Mining, Prediction Sale</i>	This study focus on implementation method <i>K-Nearest Neighbor</i> For predict sale food light at Indah Jaya Store. With use method this, with $K=3$ , the result show that algorithm <i>K-Nearest Neighbor</i> capable project amount sale food light for 6 months to front, especially for skin buffalo, with total sales reached 585 units. Research This limited to data for two years, from 2021 to 2022. The evaluation results with Root Mean Square Error (RMSE) resulting in value that is not the same with zero. More continue, focus study only on prediction sale food light and not involving prediction sale food heavy or type product other.
2	Implementation <i>Data Mining</i> for Predicting the Price of Red Onions in Yogyakarta Using the <i>K-Nearest Neighbor</i> Method [5]	<i>Data Mining, Prediction, K-Nearest Neighbor.</i>	Study This review use method <i>K-Nearest Neighbor</i> in predict sale price onion red in Yogyakarta. Approach the involving implementation algorithm <i>K-Nearest Neighbor</i> . For analyze four condition testing. Accuracy results the highest obtained was 91.67% and resulted in mark neighbor closest with produce level accuracy highest namely $K=5$ . In testing second of the 4 conditions applied to the data model. Research This only performed on onion data red, so that No can generalized for type vegetables or material food others. Therefore that, is needed study addition using a larger dataset comprehensive and able enforced in a way general.
3	Application of Predictive Data Mining Best Selling Furniture Using the <i>K-Nearest Neighbor</i> (KNN) Method [4]	<i>Data Mining, K-Nearest Neighbor, Prediction</i>	This study review use <i>K-nearest neighbor</i> method in predict sale best at Zerita Furniture Store. Through Use of <i>Data Mining</i> with technique classification and <i>K-Nearest Neighbor</i> algorithm, the results obtained is prediction that cupboard clothes become product furniture best seller, with sale reaching 90 units. Projection sale product furniture for 2020 processed through the <i>RapidMiner</i> platform. Research This only using 1 year data final for predicted so that the variables obtained only A little.
4	Implementation <i>Data Mining</i> for Prediction Sale Animal Food Products Using the <i>K-Nearest Neighbor</i> Method [3]	<i>K-Nearest Neighbor, Euclidean Distance</i>	The study utilise <i>Data Mining Techniques</i> for analyze sale best. Research This use the initial data set consists of from 360 entries sale feed animals, but after stage data cleansing, the amount shrink to 208 entries the done data sharing using <i>Data presentation</i> divided with 80% ratio for training and 20% for testing, and calculations done manually using <i>RapidMiner</i> . After set $K$ value =5, calculation done using <i>Euclidean Distance</i> , and the result is sorted based on distance closest for classification category product. Product Then grouped based on classification those. into 6 products that are less best-selling and 35 best- selling products, with level

			accuracy reached 80.4%. Research This No too Lots variation predicted goods as well as No existence information data collection during How many year final.
5	Implementation <i>Data Mining for Prediction Best Selling Frame Sales Using the K-Nearest Neighbor Method</i> [2]	Prediction <i>Sales, Data Mining, K-Nearest Neighbor</i>	Study This review implementation K-Nearest Neighbor method for predict sale frame. Approach This involving analysis large sales dataset using Data Mining. With use K-Nearest Neighbor technique in the prediction process, obtained accuracy by 88.89% with sales dataset complete and 80.0% with deleted datasets. Study This only predict frame No predict various type type wood.

Study previously use method *K-Nearest Neighbor* and applications RapidMiner *Data Mining*. Research previously only using 3 variables that is Name goods, prices and categories for predict goods electronic best seller. Difference with research that will be done that is use application *Orange*, the *tool* use between advance *drag-and-drop* so that make it easier for make channel work, data analysis, and own diverse component (*widgets*) which include data processing, modeling predictive, visualization and analysis text. For variables used researcher use a number of variable like type goods, name *brand*, price goods, units goods, stock goods, specifications products, features additional, warranty product, year product marketed and categories goods. For predict product electronic best seller at MK Store so the results obtained more accurate.

## 2.2. Data Mining Concept

According to Mustika and friends (2021) in [7] *Data Mining* is core steps in knowledge invention in data set. It integrates all analysis required procedures for disclose matter new and relevant information for interested users. Dataset possible obtained from various sources, including: traditional, databases, relational, data warehouses, web documents, or text files local simple. According to Daniel T. Larose (2004) in [7] *Data Mining* is the process of discovering valuable relationships, patterns, and trends with analyzing large stored datasets in a repository, utilizing technology introduction pattern as well as method statistics and mathematics. *Data Mining* is the process of sorting through data in amount big saved in repository for find patterns, trends, and information valuable. This technique use technology introduction patterns, statistics, mathematics, intelligence artificial, and machine learning for identify useful knowledge from various big data sources, including traditional databases, relational databases, data warehouses, web documents, or text files local.

## 2.3. Knowledge Discovery in Databases (KDD)

According to [8] *Knowledge Discovery in Database (KDD)* is extraction process series or introduction potential patterns , knowledge and information from large datasets. The knowledge and information revealed through KDD it is expected own validity, novelty, convenience understanding, and useful values. KDD is a a broad and inclusive approach various steps, start from data processing up to interpretation results. The KDD process involves steps like data selection, preprocessing, transformation, modeling, validation, and interpretation. The main objectives from KDD is For extract useful and beneficial knowledge from the existing data. The following is stages of the knowledge discovery process *in database*:



Figure 1 Knowledge Discovery in Databases (KDD)

1. **Data**  
[8]First step in the process of discovery knowledge in database (KDD) is prepare the data. The data used usually has separated from operational data.
2. **Selection**  
[8]Not all data is available can utilized with effective. so from That, a data selection process is required. The activity data selection involves target dataset formation, identification relevant variables, selection data samples, and storing data in a suitable format.
3. **Pre-processing/Cleaning**  
[8]At this stage This, the data that has been chosen will through the cleaning process. Cleaning process covers steps like delete data duplication, fix data mismatch, and correct error in the data. In addition, in stage pre-processing can also be done done data enhancement with add information relevant additions, known with term *enrichment*.
4. **Transformation**  
[8]In data mining, there are diverse algorithm, method, or available techniques. However, each algorithm, method, or technique the requires a different data format. So from that, the data that has been prepared for *the knowledge discovery in database (KDD)* process must be changed moreover first to make it fit with algorithm, method, or the technique that will used in data mining.
5. **Data Mining**  
[8]Data mining is indeed be the core of *knowledge discovery in database (KDD)*. This is a steps in which we explore and discover knowledge as well as valuable information from data using various algorithm, method, or appropriate technique with objective search desired information.
6. **Interpretation/Evaluation**  
[8]Knowledge results or information obtained from the data mining process is often presented in an easy format understood by stakeholders interests, such as graph, tree decision, or rules. Information the Then checked for verify its consistency with fact or hypothesis that has been There is previously.
7. **Knowledge**  
[8]Main objective from the *Knowledge Discovery in Databases (KDD)* process is for get knowledge or useful information from the data and implement it in accordance with benefit or desired use. This means No only about find pattern or connection in data, but also about apply knowledge found for objective certain, such as taking decisions, process improvements, or future predictions.

## 2.4. Prediction

According to Dudih Gustin and Samsul Fahmi (2020) in [7] Prediction is involving model making for predict value that has characteristics certain . The process involving estimate mark prediction based on existing patterns in data set. Prediction use a number of variable or data fields for estimate values relevant future variables.

### 3.0 METHODOLOGY

#### 3.1. KNN method

According to [9] K-Nearest Neighbor (KNN) method is approach that uses algorithm supervision For classify new test sample based on majority category from K neighbor nearest. The goal is for classify object new based on attributes and samples existing training. Algorithm This No build a special and unique model depends on the data available in memory. When given test point, algorithm search for K objects (points) train) the closest with the test point. Classification done with choose majority classification from K objects mentioned. KNN uses classification based on neighbor closest for predict mark from new test sample. The distance between neighbor usually counted use metric Euclidean distance. Here This is formula from K-Nearest Neighbor (KNN):

$$d(x, y) = \sqrt{\sum_{i=1}^n (x_i - y_i)^2} \quad (\dots \dots 1)$$

Information:

d = Proximity Distance

x = Training Data

y = Testing Data

n = Number attribute individual between 1 to with n

i = Attribute individual between 1 to with n

According to Larose (2006) in [9] For measure distance from attributes that have mark big, like attribute income, normalization often required. Normalization generally done with use min-max normalization or Z-score standardization method. If the training data consists of from mixture attribute numeric and categories, more recommended for use min-max *normalization* method. For count similarity case, used formula as following:

$$\text{Similarity } (p, q) = \frac{\sum_{i=1}^n f(p_i, q_i) \times w_i}{\sum w_i} \quad (\dots \dots 2)$$

Information :

P = New Cases

q = Existing cases in storage = Amount attribute in each case i = Attribute individual between i until with nf = Function *similarity* attribute i between case p and case a qw = Weight given to the attribute i-th . [9]

#### 3.2. Validity and Reliability Test

Before do research data analysis This conduct validity and reliability tests moreover before, for measure prediction sale product electronic best seller.

According to Sugiono (2016) in [10] Validi tool gauge refers to the extent to which the tool the produce accurate and appropriate data with what is measured. For test validity tool measuring, done with count correlation between value obtained from each statement item with mark overall from tool gauge In the research This is the method used is the Bivariate Pearson-Product Moment, the formula for which is as following :

$$r_{ij} = \frac{n \sum ix - (\sum i)(\sum x)}{\sqrt{[n \sum i^2 - (\sum i)^2][n \sum x^2 - (\sum x)^2]}} \quad (\dots \dots 5)$$

Information:

$r_{ix}$  = Coefficient item-total correlation (*bivariate person*)

i = Item score

r = Total score

n = Number of subjects taking part test

Criteria testing *Bivariate Person (Product Moment)*

1. If the calculated r value more-big from r table value (in a two- sided test) with level significance 0.05), then the instrument or statement items the own significant correlation to total score and can considered valid.

2. If the calculated r value more small from r table value (in a two- sided test) with level significance 0.05), then the instrument or statement items the No own significant correlation to total score and can considered invalid .

According to Sugiono (2016) in [10] Reliability is size reliability or data consistency from a research , which refers to the extent to which a instrument measurement can trusted For give consistent results in different times. A instrument measurement it is said reliable If give same result when used in a way repeatability. Reliability test instrument done for know consistency tool measure, namely whether tool measuring the still consistent in give same result If used in a way repeated. The reliability test method used in study This is use Alpha formula (Cronbach). Alpha formula (Cronbach) is procedure for determine mark unreliability requires item division to in two groups (although can applied to engineering split in two), so that suitable used on instruments that have odd number of items. Here is Alpha formula (Cronbach):

$$r_{ac} = \left[ \frac{k}{k-1} \right] \left[ 1 - \frac{\sum ab^2}{\sigma_{t2}} \right] \quad (\dots \dots 6)$$

Information :

- $r_{ac}$  = Reliability instrument
- $k$  = Amount grain question
- $\sum ab^2$  = Amount Variants grain
- $\sigma_{t2}$  = Total variance

Criteria evaluation Alpha (Cronbach)

Significance test done at the level significance of 0.05 means the instrument can said to be reliable if mark Alpha more big from r critical product moment.

### 3.3. Variables Used

Following is variables to be used Then tested validity and reliability. As following :

Table 2 Variables Used

No	Variables Used
1	Name of goods
2	Electrical Energy Required
3	Price of goods
4	Unit of Goods
5	Stock of goods
6	Item Weight
7	Additional Features
8	Warranty Product
9	Year Product On the market
10	Category (Target)

The table above is table from variables to be tested validity and reliability. After validity and reliability tests were conducted so variables to be used that is variables that are declared valid and reliable.

## 4.0 RESULTANTS

### 4.1. Validity and Reliability Test

Data used in validity and reliability testing is the result data questionnaire distributed to MK Store customers with total of 150 questionnaires. The questionnaires distributed through *offline*, that is with share questionnaire shaped physique paper which then distributed to Store customers. Of the 150 questionnaires distributed has get response and can collected in a way overall.

#### 4.1.1 Validity Test

Validity test questionnaire done with objective test valid or whether or not questionnaire that has been obtained in research. Testing done with to correlate results score from each criteria with score the total Then processed with use formula *Product Moment* with using 150

result data from distribution questionnaire. The results of questionnaire that has been filled in by respondents then input into excel. After That to be continued with validity and reliability tests carried out use excel.

Table 1 Validity Test

No	Criteria	r Count	r Table	Information
1	Name of goods	0.508	0.160	Valid
2	Electrical Energy Used	0.573	0.160	Valid
3	Price of goods	0.473	0.160	Valid
4	Unit of Goods	0.487	0.160	Valid
5	Stock of goods	0.599	0.160	Valid
6	Item Weight	0.613	0.160	Valid
7	Additional Features	0.458	0.160	Valid
8	Warranty Product	0.354	0.160	Valid
9	Year Product On the market	0.437	0.160	Valid
10	Item Category	0.457	0.160	Valid

Source : Writer

The table above is results from the calculated validity test use excel. Declared valid because r value calculation more big from r table value.

#### 4.1.2 Reliability Test

Reliability test done for measure consistency questionnaire which is indicator from construct or variables Questionnaire it is said reliable if answer to the question consistent from time to time. Here is results from the reliability test, as following :

Table 2 Reliability Test

No	Criteria	r ac	Information
1	Name of goods	0.661	Reliable
2	Electrical Energy Used		
3	Price of goods		
4	Unit of Goods		
5	Stock of goods		
6	Item Weight		
7	Additional Features		
8	Warranty Product		
9	Year Product On the market		
10	Item Category		

Source : Writer

The table above is results from the calculated reliability test using excel has stated reliable, with mark instrument reliability 0.661 so stated reliable. Based on the results of the validity and reliability tests that have been done so of 10 criteria on said to be valid and reliable. The following is calculation with use reliability test formula, as following :

$$r_{ac} = \left[ \frac{k}{k-1} \right] \left[ 1 - \frac{\sum ab^2}{\sigma t^2} \right]$$

$$r_{ac} = \left[ \frac{10}{10-1} \right] \left[ 1 - \frac{3,9427}{9,7496} \right]$$

$$r_{ac} = [1,111][0,5956]$$

$$r_{ac} = 0,6617$$

Calculation on counted use reliable reliability test formula. Can stated reliable Because the value of the instrument is 0.6617 then can stated reliable.

#### 4.2. K-Nearest Neighbor Method Test

Data used in data processing test with method *K-Nearest Neighbor* is results questionnaire that has been distributed to MK Store customers with total of 200 questionnaires. There are 1000 customer data in MK Store then taken 20% so that get result 200, then from 200 samples obtained Then distributed. Questionnaire the spread out through *offline*, that is with distribute to the Store customers. From 200 questionnaires distributed has get response and can collected in a way Overall. The results of calculation use excel determine the Euclidean Distance, Rank, and K=5 so that produce Category of Non-Seller Items in the 15th data table, as following:

Table 3 Excel KNN Calculation Results

$\sum_{i=1}^n f(T_i, S_i) * w_i$	$\frac{\sum_{i=1}^n f(T_i, S_i) * w_i}{\sum_{i=1}^n w_i}$	Rank	K=5
2.66	0.648780488	14	
3.1	0.756097561	9	
3.42	0.834146341	5	L
2.98	0.726829268	12	
3.74	0.912195122	2	TL
3.02	0.736585366	11	
3.74	0.912195122	1	L
3.14	0.765853659	7	
3.54	0.863414634	3	TL
3.14	0.765853659	8	
3.54	0.863414634	3	TL
2.82	0.687804878	13	
3.06	0.746341463	10	
3.38	0.824390244	6	
			TL

Source : Writer

The table above is calculation use excel, generate category in the 15th data, namely **Not Selling Well** With use mark Neighborhood closest namely K=5. Selection K value = 5 because based on *cross-validation* K=5 gives performance best and helpful reduce impact noise (data that is not relevant) and *outliers* (significant data points).

Following This is calculation with use KNN formula, as following:

$$\text{Similarity}(T, S) = \frac{\sum_{i=1}^n f(T_i, S_i) * w_i}{\sum_{i=1}^n w_i}$$

$$\text{Similarity}(T, 7) = \frac{(a1*a2)+(b1*b2)+(c1*c2)+(d1*d2)+(e1*e2)+(f1*f2)+(g1*g2)+(h1*h2)}{(a2+b2+c2+d2+e2+f2+g2+h2)}$$

$$= \frac{(1*0,8)+(0,6*0,2)+(0,6*0,2)+(1*0,4)+(0,8*1)+(1*0,5)+(1*0,6)+(1*0,4)}{(0,8+0,2+0,2+0,4+1+0,5+0,6+0,4)} = \frac{3,74}{4,1} = 0,912$$

Calculation on is a data processing test that is calculated use formula *K-Nearest Neighbor*. Can seen in the calculation use excel on T 7 produces same value which is 0.912.

#### 4.3. Orange Test

##### 4.3.1 Orange Workflow

Use *Orange* in study this is very evident effective in assist the analysis process prediction sale product electronic best seller use KNN method and *tools Orange*. Next This is channel Work wig *Orange* data mining as following:

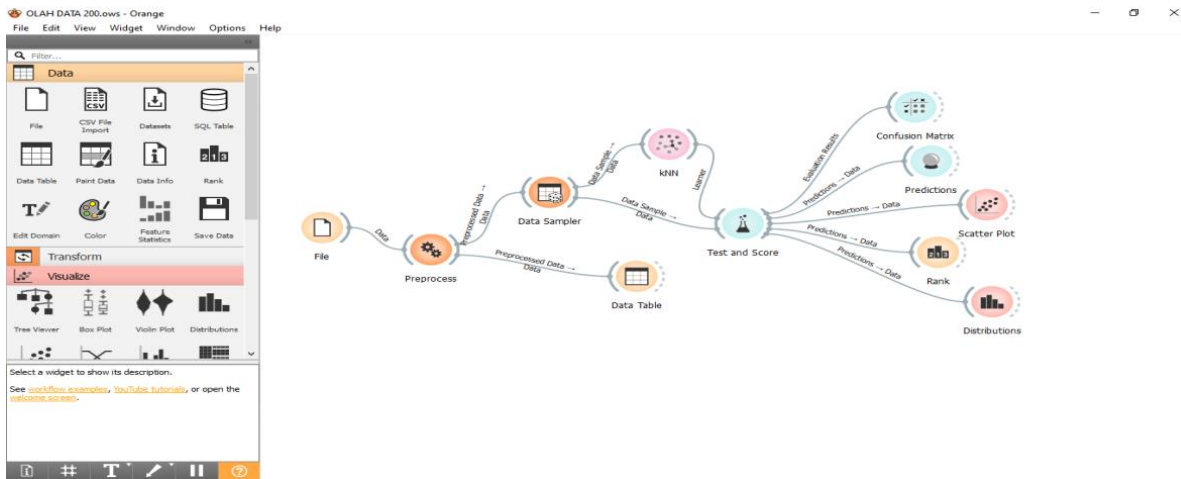


Figure 2 Orange Workflow

The image above is appearance channel work on device soft Orange. There is a number of widget used in make channel work/workflow.

### 4.3.2 Orange Distribution

Distributions on Orange functioning for display distribution mark data features in form graph. Here This is picture from Distributions, as following:

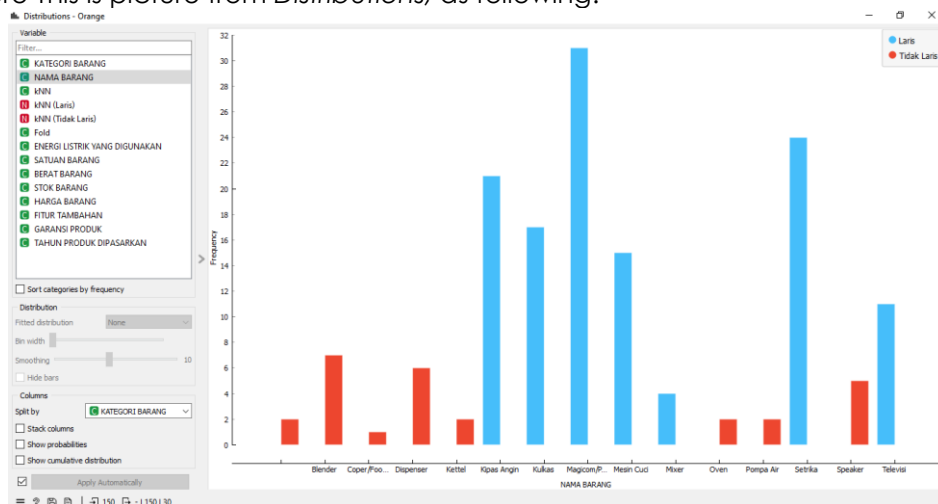


Figure 3 Distributions Orange

The image above is Distributions widget on Orange that works for display distribution in form graph. In the picture on show results that product electronic best seller at MK Store namely Magicom / Rice Cooker.

### 4.4. Results Analysis

After conduct validity and reliability tests with use excel and SPSS (Statistics Program in Social Sciences), results sale product electronic best seller with using a model or method K-Nearest Neighbor shows results that Magicom / Rice Cooker that occupies ranking highest so can it is said as product The best-selling electronics at MK Store are: Magicom / Rice Cooker. Magicom / Rice Cooker is one of the essential tools used for activity cook everyday in House stairs. Magicom used for cooking rice which is food the main thing in this country of Indonesia, then from That enthusiast goods this is very high, besides used for cooking rice function other from Magicom This is as steaming, making soup, even can make cake so from That Magicom can it is said as Product multifunctional electronics. Then availability in various size Magicom which can become consumer choice in accordance with their respective needs start from

from capacity size small up to the big one. In use *Magicom* Enough easy only with press One knob for start Cook.

*Magicom/Rice* Cooker is often the choice main in the shopping list by MK Store customers. *Magicom / Rice* Cooker becomes product electronic best seller Because his ability for fulfil need daily in House ladder with efficient. *Magicom / Functional Rice* Cooker as well as an affordable price make *Magicom* as choice main MK Store customers in Category sale product electronic best seller. The following is table from results value of AUC (*Accuracy*), CA (*Classification Accuracy*), F1 (*F1 Score*), Prec (*Precision*), Recall (*True Positive Rate*), and MCC (*Matthews Correlation Coefficient*), as following :

Table 4 Accuracy Values

Model	AUC	CA	F1	Prec	Recall	MCC
<i>K-Nearest Neighbor</i>	0.962	0.980	0.980	0.980	0.980	0.932

From the table on can concluded that AUC (*Accuracy*) level method algorithm *K-Nearest Neighbor* is very high that is reached 0.962 or in mark presentation which is 96%. The CA (*Classification Accuracy*) value level reached 0.980, level F1 value (*F1 Score*) reached 0.980, the level mark Prec (*Precision*) 0.980, level Recall value (*True Positive Rate*) reached 0.980, and the level MCC value (*Matthews Correlation Coefficient*) reached 0.932. In the study previously conducted by [11] produce mark accuracy reached 96.15%, the level mark Precc reached 94.20%, the level Recall value reached 100%, and the level Specificity value reached 89.74%. In the research conducted by [2, p. 12] produce mark accuracy 88.89%, resulting in level Recall value 100% and produces level mark Prec by 95%. In the research conducted by [3, p. 4] produce level mark accuracy of 80.4%, resulting in level mark Prec 88%, and produces level Recall value 88.5%. In the research conducted by [16, p. 5] produce level mark accuracy reaching 80%, resulting in mark Prec 80%, yield level F1 value 89%, and produces Recall value 100%. In the research conducted by [17, p. 6] produce level mark accuracy reached 93.35%, resulting in level mark Prec 88.14%, and produces level Recall value 100%. From several study previous produce level mark high accuracy and precision with use same method that is *K-Nearest Neighbor* so from That can proven that with using algorithmic models *K-Nearest Neighbor* proven its accuracy.

## 5.0 CONCLUSION

Based on discussion that has been outlined about predict sale product electronic best seller at MK Store using algorithm *K-Nearest Neighbor*, then can concluded that model or function that describes *class* sale product electronic best seller use a number of criteria that is including Item Name, Electrical Energy Used, Item Price, Item Unit, Item Stock, Item Weight, Additional Features, Warranty Product, Year Product Marketed and Product Category. From several criteria which is used as base from measurement Then processed use *orange tools* for build a measurement model that is described use mark Neighborhood closest, then for know goods electronic best seller with use *wig distributions* that display picture in the form of chart for to know product the best-selling electronics that produce product *Magicom / Rice* Cooker. Here is order from start from the highest until lowest chart *wighet distributions* include *Magicom/Rice* Cooker, Setika, Fan Wind, Refrigerator, Machine Washing, Television, Mixer, Blender, Dispenser, Speaker, Oven, Coper, Kettel, Water Pump and Air Frayer. Then For feasibility of the obtained model supported with level accuracy and precision from algorithm tester. The AUC (*Accuracy*) value level is 0.962, the level CA (*Classification Accuracy*) value reached 0.980, the level F1 value (*F1 Score*) reached 0.980, the level mark Prec (*Precision*) 0.980, level Recall value (*True Positive Rate*) reached 0.980, and the level MCC value (*Matthews Correlation Coefficient*) reached 0.932.

## Acknowledgement

Writer to pronounce accept love to **DRTPM Ministry of Education, Culture, Research and Technology** which has give support **financial on Research of New Lecturers (PDP) Affirmation year 2024 Funding** and Speech Thank You to Chairman of the Bhima Sakti Lampung Foundation and Chancellor Bakti Nusantara Institute which has give input and direction on the research conducted.

## REFERENCES

- [1] Dataindonesia.id, "Data Badan Pusat Statistik," *dataindonesia.id*, 2023. .
- [2] R. Bahtiar, "Implementasi Data Mining Untuk Prediksi Penjualan Kusen Terlaris Menggunakan Metode K-Nearest Neighbor," *J. Inform. MULTI*, vol. 1, no. 3, pp. 200–214, 2023.
- [3] R. Iriane and Nurfaizah, "Penerapan Data Mining Untuk Prediksi Penjualan Produk Pangan Hewan Menggunakan Metode K-Nearest Neighbor," *Media Online*, vol. 3, no. 5, pp. 509–515, 2023.
- [4] N. E. Pratiwi, L. Suryadi, N. E. Pratiwi, F. Ardhy, and P. Riswanto, "JUSIM ( Jurnal Sistem Informasi Musirawas ) Penerapan Data Mining Prediksi Penjualan Mebel Terlaris Menggunakan Metode K-Nearest Neighbor (K-NN) (Studi Kasus: Toko Zerita Meubel)," *JUSIM (Jurnal Sist. Inf. Musirawas) Vol. 7, No. 2, Desember 2022*, vol. 7, no. 2, pp. 174–184, 2022.
- [5] D. Virdaus and P. T. Prasetyaningrum, "Penerapan Data Mining Untuk Memprediksi Harga Bawang Merah Di Yogyakarta Menggunakan Metode K-Nearest Neighbor," *J. ....*, no. 84, pp. 1–8, 2021.
- [6] Y. Handayani, T. Hidayat, and H. Arruhama, "Implementasi Metode K-Nearest Neighbor Untuk Prediksi Penjualan Produk Terlaris Pada Toko Indah Jaya," vol. 2, pp. 82–91, 2023.
- [7] dan S. I. Rulin Swastika, Siti Mukodimah, Ferry Susanto, Muhamad Muslihudin, *Implementasi Data Mining (Clustering, Association, Prediction, Estimation, Classification)*, 1st ed. Indramayu: CV. Adanu Abimata, 2023.
- [8] E. Buulolo, *Data Mining Untuk Perguruan Tinggi*. Deepublish, 2020.
- [9] Yahya, *Data Mining*. CV Jejak, anggota IKAPI, 2020.
- [10] S. Mukodimah, "Implementasi Algoritma C4.5 Dan Naive Bayes Untuk Mengukur Efektivitas Pembelajaran Daring Di Wilayah Pringsewu," 2022.
- [11] Rismala, I. Ali, and A. R. Rinaldi, "Penerapan Metode K-Nearest Neighbor Untuk Prediksi Penjualan Sepeda Motor Terlaris," vol. 7, no. 1, pp. 585–590, 2023.
- [12] M. I. Hutapea and A. P. Silalahi, "Moderna's Vaccine Using the K-Nearest Neighbor (KNN) Method: An Analysis of Community Sentiment on Twitter," *J. Penelit. Pendidik. IPA*, vol. 9, no. 5, pp. 3808–3814, 2023.
- [13] A. Qurotul and M. Reza, "Application Of The K-Nearest Neighbor Algorithm For Student Department Classification At 15 Pekanbaru State High School Penerapan Algoritma K-Nearest Neighbor Untuk Klasifikasi Jurusan Siswa Di Sma Negeri 15 Pekanbaru," vol. 3, no. 1, pp. 39–45, 2023.