



# Plagiarism Checker X Originality Report

**Similarity Found: 13%**

Date: Senin, Juli 27, 2020

Statistics: 253 words Plagiarized / 1879 Total words

Remarks: Low Plagiarism Detected - Your Document needs Optional Improvement.

---

Analisis Matthew **Correlation Coefficient** pada **K-Nearest Neighbor** dalam Klasifikasi **Ikan Hias** Novia Hasdyna<sup>1</sup>, Rozzi Kesuma dinata<sup>2</sup> 1Program Studi Teknik Informatika, Universitas Islam Kebangsaan Indonesia 2,Program Studi Teknik Informatika, Universitas Malikussaleh 1noviahasdyna@gmail.com, 2rozzi@unimal.ac.id ABSTRACT \_ \_K-Nearest Neighbor (K-NN) is a **machine learning** algorithm that functions **to classify** data.

This study aims **to measure the performance of** K-NN algorithm by using Matthew **Correlation Coefficient** (MCC). **The data** that used **in this** study **are the** ornamental fish which consisting of 3 classes named Premium, Medium, and Low. The analysis results **of the** Matthew **Correlation Coefficient** on K-NN using **Euclidean Distance** obtained the highest MCC value in Medium class which is 0.786542.

The second highest MCC value is in Premium class which is 0.567434. The lowest MCC value is in Low class which is 0.435269. Overall, the MCC values is statistically which is 0,596415. \_ \_ Keyword: Matthew correlation coefficient, **machine learning**, knn, **ikan hias** \_ \_ Introduction Saat ini perkembangan ruang lingkup **analisis data** science yang sedang **berkembang pesat** adalah machine learning [1].

Machine learning atau **pembelajaran mesin** merupakan bagian dari artificial intelligence yang **terdiri dari supervised learning, unsupervised learning, dan reinforcement learning** [2]. Algoritma **supervised learning** antara lain adalah support vector machine, naïve bayes, k-nearest neighbor, dan **random forest** [3]. **Supervised learning** berfungsi **untuk mengklasifikasikan data berdasarkan pola yang sudah ada** [4].

Banyak **metode yang** dipakai untuk mengukur kinerja algoritma **machine learning** [5]. Pada penelitian ini, **untuk mengukur kinerja** algoritma K-Nearest Neighbor digunakan

matthew correlation coefficient. Matthew correlation coefficient mengukur performansi klasifikasi data dengan range -1, 0, +1 [6].

Semakin nilai mcc mendekati maka semakin baik kinerja algoritma klasifikasinya. Sebaliknya, jika nilai nya mendekati -1, maka semakin buruk kinerja algoritma klasifikasi. Pengukuran terhadap kinerja suatu sistem klasifikasi merupakan hal yang penting [7]. Kinerja sistem klasifikasi menggambarkan seberapa baik sistem dalam mengklasifikasikan data.

Penelitian ini menggunakan algoritma k-nn untuk mengklasifikasikan data ikan hias. Adapun klasifikasi ikan hias terdiri dari 3 class, yaitu premium, medium dan low. Hasil yang diharapkan pada penelitian ini berupa nilai performance Matthew Correlation Coefficient (MCC) pada algoritma knn dalam mengklasifikasikan data ikan hias. Pada proses klasifikasi ikan hias ini akan memberikan informasi hasil pengujian data ikan hias termasuk kedalam grade premium, medium atau grade low berdasarkan konversi nilai pada atribut perawatan, warna, pakan, dan harga. Research Method 2.1.

Framework Penelitian Adapun framework penelitian ini adalah seperti pada gambar 1. Gambar 1. Framework Penelitian Berdasarkan gambar 1, skema penelitian ini terdiri dari tiga proses yaitu input, proses dan output. Setelah melakukan inputan data ikan hias, selanjutnya akan dilakukan proses perhitungan jarak dengan menggunakan Euclidean distance.

Jarak terkecil akan dilakukan proses pengklasifikasian dengan k-nearest neighbor. Hasil klasifikasi berdasarkan data training dan data testing akan dianalisis performance nya dengan menggunakan matthew correlation coefficient. Semakin nilai mcc mendekati maka semakin baik kinerja algoritma klasifikasinya. Sebaliknya, jika nilai nya mendekati -1, maka semakin buruk kinerja algoritma klasifikasi.

2.2. K-Nearest Neighbor Salah satu metode klasifikasi terhadap sekumpulan data berdasarkan pembelajaran data yang sudah terklasifikasikan sebelumnya adalah K-NN [8]. K-NN termasuk dalam golongan supervised learning, dimana hasil query instance yang baru diklasifikasikan berdasarkan mayoritas kedekatan jarak dari kategori yang ada dalam K-NN [9]. Berikut ilustrasi K-NN seperti pada gambar 2. Gambar 2.

Ilustrasi K-NN dengan k=1, k=2, dan k=3 K- Nearest Neighbor bekerja mencari jarak yang paling dekat antara data yang akan di evaluasi dengan k neighbor ( tetangga ) yang terdekat di dalam sebuah data training [10]. Berikut urutan proses kerja algoritma K- Nearest Neighbor [11]: Tentukan Parameter k jumlah tetangga paling dekat. Hitung Euclidean Distance masing masing objek terhadap sampel data yang ada.

(1) Keterangan:  $x_1$  = Sampel data  $x_2$  = Data uji atau data testing  $i$  = Variabel data  $d$  = Jarak  $p$  = Dimensi data Kemudian mengurutkan objek-objek tersebut kedalam kelompok yang mempunyai jarak Euclid kecil. Mengumpulkan kategori Y (Klasifikasi Nearest Neighbor). 2.3 Matthew Correlation Coefficient Matthew Correlation Coefficient (MCC) merupakan salah satu metode yang digunakan untuk mengukur kinerja algoritma klasifikasi.

Perhitungan awal mcc dilakukan dengan menggunakan confusion matrix. Pada dasarnya confusion matrix mengandung informasi yang membandingkan hasil klasifikasi yang dilakukan oleh sistem dengan hasil klasifikasi yang seharusnya. Koefisien Korelasi Matthews (MCC) memiliki rentang -1 hingga 1 di mana -1 menunjukkan klasifikasi biner yang sepenuhnya salah sedangkan 1 menunjukkan klasifikasi biner yang sebenarnya. Menggunakan MCC memungkinkan seseorang untuk mengukur seberapa baik kinerja model/fungsi klasifikasi.

Berikut adalah formula dari MCC: Keterangan: FN = False Negative TP = True Positive TN = True Negative FP = False Positive MCC = Matthew Correlation Coefficient Result and Analysis 3.1 Data Training Adapun dataset ikan hias dapat dilihat pada tabel 1.

Tabel 1. Data Training Nama Ikan \_Perawatan \_Warna \_Pakan \_Harga \_Grade \_ \_Fish A \_5 \_Sangat Menarik \_Binatang Kecil \_Puluhan Juta \_Premium \_ \_Fish B \_5 \_Sangat Menarik \_Binatang Kecil dan Pelet \_Puluhan Juta \_Premium \_ \_Fish C \_5 \_Sangat Menarik \_Pelet dan Tumbuhan \_Jutaan \_Premium \_ \_Fish D \_4 \_Sangat Menarik \_Pelet dan Tumbuhan \_Jutaan \_Premium \_ \_Fish E \_5 \_Sangat Menarik \_Pelet \_Ratusan Ribu \_Premium \_ \_Fish F \_4 \_Menarik \_Binatang Kecil dan Pelet \_Puluhan Ribu \_Medium \_ \_Fish G \_3 \_Sangat Menarik \_Pelet dan Tumbuhan \_Puluhan Ribu \_Medium \_ \_Fish H \_3 \_Sangat Menarik \_Pelet \_Puluhan Ribu \_Medium \_ \_Fish I \_3 \_Menarik \_Pelet \_Puluhan Ribu \_Medium \_ \_Fish J \_3 \_Menarik \_Tumbuhan \_Puluhan Ribu \_Medium \_ \_Fish K \_2 \_Menarik \_Pelet \_Ribuan \_Low \_ \_Fish L \_2 \_Netral \_Pelet \_Ribuan \_Low \_ \_Fish M \_2 \_Kurang Menarik \_Pelet \_Ribuan \_Low \_ \_Fish N \_2 \_Sangat Kurang Menarik \_Pelet \_Ribuan \_Low \_ \_Fish O \_1 \_Kurang Menarik \_Tumbuhan \_Ribuan \_Low \_ \_ Tabel 2.

Atribut dan Nilainya No \_Nama \_Jenis atribut dan Nilainya \_ \_1 \_Ikan Hias \_Kategori (Nama Ikan) \_ \_2 \_Kemudahan Perawatan \_Numerik \_ \_3 \_Warna \_Kategori (Sangat Menarik, Menarik, Netral, Kurang Menarik, Sangat Kurang Menarik) \_ \_4 \_Pakan \_Kategori (Binatang Kecil dan pelet, Pelet dan Tumbuhan, Binatang Kecil, Tumbuhan, Pelet) \_ \_5 \_Harga \_Kategori (Puluhan Juta, Jutaan, Ratusan ribu, Puluhan ribu, Ribuan) \_ \_ 3.2. Data Testing Tabel 3.

Data Testing Nama Ikan \_Perawatan \_Warna \_Pakan \_Harga \_Grade \_ \_Blue Tang \_5 \_4 \_2

\_3 ? \_ Pare Tawar \_4 \_2 \_5 \_3 ? \_ Clown Fish \_4 \_4 \_2 \_2 ? \_ Gurame \_3 \_4 \_1 \_1 ? \_  
\_Blue Tang \_5 \_4 \_2 \_3 ? \_ **Tabel 4.** Konversi Warna Kriteria Warna \_Nilai \_ \_Sangat  
Menarik \_5 \_ \_Menarik \_4 \_ \_Netral \_3 \_ \_**Kurang Menarik** \_2 \_ \_Sangat Kurang Menarik \_1  
\_ \_ **Tabel 5.**

Konversi Pakan Kriteria Pakan \_Nilai \_ \_Binatang Kecil \_5 \_ \_Binatang Kecil dan Pelet \_4 \_  
\_Pelet dan Tumbuhan \_3 \_ \_Pelet \_2 \_ \_Tumbuhan \_1 \_ \_ **Tabel 6.** Konversi Harga Kriteria  
Harga \_Nilai \_ \_Puluhan Juta \_5 \_ \_Jutaan \_4 \_ \_Ratusan Ribu \_3 \_ \_Puluhan Ribu \_2 \_  
\_Ribuan \_1 \_ \_ **Tabel 7.** Konversi **Data Training** Nama Ikan \_Perawatan \_Warna \_Pakan  
\_Harga \_Grade \_  
\_Fish A \_5 \_5 \_5 \_5 \_Premium \_ \_Fish B \_5 \_5 \_4 \_5 \_Premium  
\_ \_Fish C \_5 \_5 \_3 \_4 \_Premium \_ \_Fish D \_4 \_5 \_3 \_4 \_Premium \_ \_Fish E \_5 \_5 \_2 \_3  
\_Premium \_ \_Fish F \_4 \_4 \_4 \_2 \_Medium \_ \_Fish G \_3 \_5 \_3 \_2 \_Medium \_ \_Fish H \_3 \_5 \_2  
\_2 \_Medium \_ \_Fish I \_3 \_4 \_2 \_2 \_Medium \_ \_Fish J \_3 \_4 \_1 \_2 \_Medium \_ \_Fish K \_2 \_4 \_2  
\_1 \_Low \_ \_Fish L \_2 \_3 \_2 \_1 \_Low \_ \_Fish M \_2 \_2 \_2 \_1 \_Low \_ \_Fish N \_2 \_1 \_2 \_1 \_low \_  
\_Fish O \_1 \_1 \_1 \_1 \_Low \_ \_ **3.3 Hasil Klasifikasi K-NN 3.3.1. Hasil Perhitungan Jarak**  
**dengan Euclidean Distance** **Tabel 8.**

Hasil **Perhitungan Jarak** Data Test 1 (Blue Tang) Nama Ikan \_Jarak \_Grade \_ \_ \_ \_ \_ \_ \_ \_ \_ \_  
\_Fish A \_3,74 \_Premium \_ \_Fish B \_3,00 \_Premium \_ \_Fish C \_1,73 \_Premium \_ \_Fish D \_2,00  
\_Premium \_ \_Fish E \_1,00 \_Premium \_ \_Fish F \_2,45 \_Medium \_ \_Fish G \_2,65 \_Medium \_  
\_Fish H \_2,45 \_Medium \_ \_Fish I \_2,24 \_Medium \_ \_Fish J \_2,45 \_Medium \_ \_Fish K \_3,61  
\_Low \_ \_Fish L \_3,74 \_Low \_ \_Fish M \_4,12 \_Low \_ \_Fish N \_4,69 \_low \_ \_Fish O \_5,48 \_Low \_  
\_ **Tabel 9.**

Hasil **Perhitungan Jarak** Data Test 2 (Pare Tawar) Nama Ikan \_Jarak \_Grade \_ \_ \_ \_ \_ \_ \_ \_ \_ \_  
\_Fish A \_3,74 \_Premium \_ \_Fish B \_3,87 \_Premium \_ \_Fish C \_3,87 \_Premium \_ \_Fish D \_3,74  
\_Premium \_ \_Fish E \_4,36 \_Premium \_ \_Fish F \_2,45 \_Medium \_ \_Fish G \_3,87 \_Medium \_  
\_Fish H \_4,47 \_Medium \_ \_Fish I \_3,87 \_Medium \_ \_Fish J \_4,69 \_Medium \_ \_Fish K \_4,58  
\_Low \_ \_Fish L \_4,24 \_Low \_ \_Fish M \_4,12 \_Low \_ \_Fish N \_4,24 \_low \_ \_Fish O \_5,48 \_Low \_  
\_ **Tabel 10.**

Hasil **Perhitungan Jarak** Data Test 3 (Clown Fish) Nama Ikan \_Jarak \_Grade \_ \_ \_ \_ \_ \_ \_ \_ \_ \_  
\_Fish A \_4,47 \_Premium \_ \_Fish B \_3,87 \_Premium \_ \_Fish C \_2,65 \_Premium \_ \_Fish D \_2,45  
\_Premium \_ \_Fish E \_1,73 \_Premium \_ \_Fish F \_2,00 \_Medium \_ \_Fish G \_1,73 \_Medium \_  
\_Fish H \_1,41 \_Medium \_ \_Fish I \_1,00 \_Medium \_ \_Fish J \_1,41 \_Medium \_ \_Fish K \_2,24  
\_Low \_ \_Fish L \_2,45 \_Low \_ \_Fish M \_3,00 \_Low \_ \_Fish N \_3,74 \_low \_ \_Fish O \_4,47 \_Low \_  
\_ **Tabel 11.**

Hasil **Perhitungan Jarak** Data Test 4 (Gurame) Nama Ikan \_Jarak \_Grade \_ \_ \_ \_ \_ \_ \_ \_ \_ \_  
\_Fish A \_6,08 \_Premium \_ \_Fish B \_5,48 \_Premium \_ \_Fish C \_4,24 \_Premium \_ \_Fish D \_3,87

\_Premium \_ \_Fish E \_3,16 \_Premium \_ \_Fish F \_3,32 \_Medium \_ \_Fish G \_2,45 \_Medium \_  
\_Fish H \_1,73 \_Medium \_ \_Fish I \_1,41 \_Medium \_ \_Fish J \_1,00 \_Medium \_ \_Fish K \_1,41  
\_Low \_ \_Fish L \_1,73 \_Low \_ \_Fish M \_2,45 \_Low \_ \_Fish N \_3,32 \_low \_ \_Fish O \_3,61 \_Low \_  
\_3.4

Hasil analisis Matthew **Correlation Coefficient** Tabel 12. **Proses I Confusion Matrix**  
predicted class \_ \_premium \_medium \_low \_ \_ \_premium \_3 \_0 \_0 \_ \_ \_medium \_0 \_4 \_1 \_ \_  
\_low \_0 \_0 \_4 \_ \_ Tabel 13. **Proses II Confusion Matrix** predicted class \_ \_TP \_FP \_FN \_TN \_  
\_ \_premium \_3 \_0 \_0 \_2 \_ \_ \_medium \_4 \_0 \_0 \_1 \_ \_ \_low \_4 \_0 \_0 \_1 \_ \_ Tabel 14.

Nilai Matthew **Correlation Coefficient** No \_class \_mcc \_ \_1 \_Premium \_0,567434 \_ \_2  
\_medium \_0,786542 \_ \_3 \_low \_0,435269 \_ \_ Conclusion **Penelitian ini** menganalisis  
kinerja algoritma klasifikasi KNN **dalam pengklasifikasian** data ikan hias. Hasil **penelitian**  
**ini** menunjukkan bahwa nilai matthew **correlation coefficient** pada k-nn **dengan**  
**menggunakan Euclidean distance** diperoleh nilai MCC tertinggi pada class medium  
sebesar 0,786542. Nilai mcc tertinggi kedua pada class premium senilai 0,567434. Nilai  
mcc terendah adalah pada class low sebesar 0,435269.

Dengan demikian, dapat dikatakan nilai korelasi matthew **yang paling** baik adalah pada  
class medium karena class medium adalah **yang paling** mendekati +1. Sebaliknya nilai  
MCC pada class low adalah **yang paling** mendekati 0. Adapun rata-rata nilai MCC pada  
**klasifikasi data ikan hias** dengan KNN adalah senilai 0,596415. **Oleh karena itu,** kinerja  
algoritma klasifikasi KNN adalah baik.

Adapun saran dari penulis adalah agar pembaca dapat melakukan analisis kinerja  
algoritma klasifikasi **dengan metode yang** lain, seperti metode k-fold cross validation.

#### INTERNET SOURCES:

1% -

[https://www.researchgate.net/publication/270163293\\_Application\\_of\\_k-Nearest\\_Neighbour\\_Classification\\_in\\_Medical\\_Data\\_Mining](https://www.researchgate.net/publication/270163293_Application_of_k-Nearest_Neighbour_Classification_in_Medical_Data_Mining)

1% - <https://link.springer.com/article/10.1186/s12859-019-2766-9>

<1% -

[https://www.cambridge.org/core/services/aop-cambridge-core/content/view/26A88C1604E046266B8C5F8F0544CE0D/S1068280518000096a.pdf/climate\\_land\\_cover\\_and\\_bird\\_populations\\_differential\\_impacts\\_on\\_the\\_future\\_welfare\\_of\\_birders\\_across\\_the\\_pacific\\_northwest.pdf](https://www.cambridge.org/core/services/aop-cambridge-core/content/view/26A88C1604E046266B8C5F8F0544CE0D/S1068280518000096a.pdf/climate_land_cover_and_bird_populations_differential_impacts_on_the_future_welfare_of_birders_across_the_pacific_northwest.pdf)

<1% -

<https://masoemuniversity.ac.id/berita/machine-learning-dan-generasi-millennia-part1.p>

hp

1% - <https://repository.unpak.ac.id/tukangna/repo/file/files-20190131070842.pdf>

1% -

<https://mc.ai/implementasi-deep-learning-menggunakan-convolutional-neural-network-untuk-klasifikasi-gambar-mata/>

<1% - <https://ojs.uniska-bjm.ac.id/index.php/JIT/article/download/1128/950>

<1% - <http://eprints.mercubuana-yogya.ac.id/955/3/BAB%20II.pdf>

1% - <https://ejournal.upm.ac.id/index.php/energy/article/download/111/449/>

<1% - <https://lordbroken.wordpress.com/category/ilmu-dan-teknologi-pangan/page/7/>

<1% - <https://repository.ugm.ac.id/view/year/2013.type.html>

1% - [http://repository.upi.edu/34313/6/S\\_KOM\\_1304373\\_Chapter3.pdf](http://repository.upi.edu/34313/6/S_KOM_1304373_Chapter3.pdf)

<1% - <http://digilib.unila.ac.id/9187/16/BAB%20II.pdf>

<1% - <https://pt.scribd.com/doc/251810302/Proceedings-Knsi-2014-Full-Edition>

1% - <http://publikasi.dinus.ac.id/index.php/technoc/article/download/1245/927>

<1% - <https://id.scribd.com/presentation/358699677/romi-dm-02-proses-nov2016-pptx>

1% - <https://conference.upnvj.ac.id/index.php/seinasikesi/article/download/73/pdf>

1% -

[https://www.researchgate.net/publication/316591760\\_IG-KNN\\_UNTUK\\_PREDIKSI\\_CUSTOMER\\_CHURN\\_TELEKOMUNIKASI](https://www.researchgate.net/publication/316591760_IG-KNN_UNTUK_PREDIKSI_CUSTOMER_CHURN_TELEKOMUNIKASI)

1% -

[http://komputa.if.unikom.ac.id/\\_s/data/jurnal/volume-01/komputa-1-1-perbandingan-algoritma-template-irfan-maliki-5.pdf/pdf/komputa-1-1-perbandingan-algoritma-template-irfan-maliki-5.pdf](http://komputa.if.unikom.ac.id/_s/data/jurnal/volume-01/komputa-1-1-perbandingan-algoritma-template-irfan-maliki-5.pdf/pdf/komputa-1-1-perbandingan-algoritma-template-irfan-maliki-5.pdf)

<1% - <https://edoc.pub/contributor/103805436/Penerapan-Algorithm-Modified-K-nearest-Neighbor-Mknn-Untuk-Mengklasifikasikan-Letak-Protein-Pa2>

<1% -

<https://id.scribd.com/doc/103805436/Penerapan-Algorithm-Modified-K-nearest-Neighbor-Mknn-Untuk-Mengklasifikasikan-Letak-Protein-Pa2>

1% - <http://repository.unmuhjember.ac.id/589/1/JURNAL.pdf>

<1% -

<https://id.123dok.com/document/zpn1wkoy-studi-kasus-evolusi-proyek-perangkat-lunak-open-source-weka.html>

1% - <http://jurnal.unsil.ac.id/index.php/innovatics/article/download/872/667>

1% -

[https://www.researchgate.net/publication/235902451\\_Data\\_Mining\\_Concept\\_and\\_Techniques](https://www.researchgate.net/publication/235902451_Data_Mining_Concept_and_Techniques)

1% - <http://ejournal.nusamandiri.ac.id/index.php/pilar/article/download/28/23/>

1% - <https://achmatim.net/tag/klasifikasi/>

<1% -

<https://www.lppm.itb.ac.id/wp-content/uploads/sites/55/2020/03/Evaluasi-Diri-Kantor-LPPM-ITB-2019.pdf>

<1% - <https://bukubiruku.com/jenis-ikan-hias-air-tawar-aquarium/>  
<1% - [https://issuu.com/waspada/docs/waspada\\_kamis\\_27\\_februari\\_2014](https://issuu.com/waspada/docs/waspada_kamis_27_februari_2014)  
<1% -  
<http://repository.uinjkt.ac.id/dspace/bitstream/123456789/9466/1/ISRAWATI-PSI.pdf>  
<1% - <http://digilib.unila.ac.id/6427/118/BAB%20III.pdf>  
<1% - <https://sitimariyam626.wordpress.com/category/uncategorized/>  
<1% -  
[https://mafiadoc.com/kumpulan-makalah-pkmp-ristek-bem-its-10-11\\_5a1888161723dd6415fd741b.html](https://mafiadoc.com/kumpulan-makalah-pkmp-ristek-bem-its-10-11_5a1888161723dd6415fd741b.html)  
<1% - <https://nepis.epa.gov/Exe/ZyPURL.cgi?Dockey=91019A8Q.TXT>  
1% - <https://journal.unnes.ac.id/nju/index.php/JIPK/article/download/10623/6483>  
1% - <https://repository.telkomuniversity.ac.id/catalogue/2011.html>  
<1% - <https://d-ys.blogspot.com/feeds/posts/default>  
1% -  
<https://repository.bsi.ac.id/index.php/unduh/item/237648/Jurnal-Khatulistiwa-genap-2019.pdf>  
<1% - <https://www.scribd.com/document/390942873/118-5-PB-pdf>