

IMPLEMENTASI ALGORITMA MODIFIED K-NEAREST NEIGHBOR (M-KNN) UNTUK KLASIFIKASI DIAGNOSA PENYAKIT PADA KUCING

Aprela Kurnia Putri¹, M. Tanzil Furqon², Mahendra Data³

Program Studi Informatika, Fakultas Ilmu Komputer
Universitas Brawijaya

Jl. Veteran, Malang 65145, Indonesia

Email: Aprelania94@gmail.com, m.tanzil.furqon@gmail.com, mahendra.data@ub.ac.id

ABSTRAK

Kucing bahkan dari sejak masih kecil adalah salah satu hewan favorit manusia untuk dipelihara dan disayangi sebagai peliharaan. Sebuah penelitian di Universitas Muhammadiyah Malang oleh Setianingrum membuktikan bahwa memelihara hewan peliharaan membantu manusia untuk melepaskan stress dengan menganggap hewan peliharaan sebagai teman diberbagai aktivitas. Dan beberapa tahun belakangan ini, orang-orang mulai lebih memperhatikan dan mulai memberikan perlakuan terbaik untuk hewan peliharaannya. Oleh karena itu, pemilik peliharaan tersebut harus memperhatikan kesehatan kucing. Untuk itu diperlukan penanganan yang cepat sehingga pemilik peliharaan dapat melakukan penanggulangan dengan segera. Namun kenyataannya, klinik yang ditangani langsung oleh dokter tidak buka selama 24 jam. Untuk mengatasi hal tersebut maka dibuatlah sebuah sistem dengan memanfaatkan salah satu metode implementasi dari data mining yaitu metode *Modified K-Nearest Neighbor* (MKNN). *Modified K-Nearest Neighbor* (MKNN) adalah pengembangan dari metode KNN yang dirancang untuk mengatasi kelemahan dari jarak data dengan *weight* pada *outlier* KNN. Terdapat 18 gejala dan 8 jenis penyakit yang dapat digunakan sebagai parameter dalam pengembangan sistem. Keluaran yang dihasilkan sistem berupa diagnosa penyakit. Berdasarkan pada skenario pengujian yang dilakukan memperoleh hasil rata-rata akurasi maksimum 99,90%, dengan *precision* sebesar 99,76% dan *recall* 99,41% serta akurasi minimum sebesar 79,76%. Berdasarkan hasil tersebut, sistem yang menggunakan metode *Modified K-Nearest Neighbor* (MKNN) dapat diimplementasikan kehidupan sehari-hari.

Kata Kunci: *Modified K-Nearest Neighbor* (M-KNN), Klasifikasi, Penyakit Kucing.

ABSTRACT

Cat, ever since it was just a kitty, is one of favorite animal to be cared and loved by human as their pet. Not only because it easy to raise, but also it can release stress. A Study in University of Muhammadiyah Malang by Setianingrum tells us that petting a pet help peoples to realise their stress by using a pet as their friend in any activities. And last few years people start to care more about their pet and start to give their pet best treatment including health. Therefore, it is necessary for quick handling so that pet owner can do countermeasures immediately. But in reality, the clinic that handled directly by the doctor is not available for 24 hours. To solve this problem a system is needed in order to satisfy costumer or pet owner. An implementation of data mining that use the *Modified k-nearest neighbor* (MKNN.) *Modified k-nearest neighbor* (MKNN) is the development of KNN method designed to overcome the disadvantages of data distance by weight on KNN outlier. There are 18 symptoms and 8 types of diseases that is used as a parameter in development of the system. An output as the result produced by the system is disease diagnosis. based on method, this research obtain 99,90% of maximum accuracy for an average yield with precision around 99,76% and recall around 99,41%, and 79,76% minimum accuracy. The System can be applied in society based on result.

Keywords: Modified K-Nearest Neighbor (M-KNN), Classification and Cat Diseases.

1. PENDAHULUAN

1.1 Latar Belakang

Salah satu hewan peliharaan yang banyak digemari untuk dipelihara adalah kucing. Kucing merupakan hewan karnivora yang memiliki banyak jenis seperti kucing Persia, Anggora, *Sphynx*, dan

lain-lain. Dalam keseharian, kebanyakan orang yang memelihara kucing rumahan menjadikannya sebagai teman bahkan kucing tersebut dapat menjadi sangat dekat dengan manusia seperti

keluarga. Bagi pemiliknya tentu saja hewan peliharaan seperti kucing ini dapat membawa pengaruh dalam kehidupan sehari-hari pemilik. Kucing dapat membawa kebahagiaan dan mengurangi rasa stress. Selain untuk dipelihara kucing dapat memberikan peluang bagi pemiliknya untuk mengembangkan kucing tersebut sehingga dapat membuka peluang dalam hal perekonomian. (Holbrook B., 2008)

Maka dari itu, banyak pemilik kucing yang ingin memberikan perhatian khusus kepada kucing peliharaannya. Utamanya adalah masalah kesehatan kucing. Terkadang penyakit yang menyerang hewan peliharaan tersebut dapat membuat pemilik merasa bingung. Hal ini dikarenakan kurangnya pengetahuan terhadap diagnosa serta penanganan penyakit yang menyerang peliharaannya maupun kesulitan menemui dokter hewan dalam keadaan tertentu. Kenyataannya ketika hewan tersebut sedang terserang penyakit, belum ada klinik yang ditangani langsung oleh dokter hewan yang buka selama 24 jam. Diperlukan penanganan yang cepat sehingga pemilik peliharaan dapat melakukan penanggulangan dengan segera. Untuk mengatasi hal tersebut, perlu adanya sebuah sistem yang mendiagnosa penyakit agar pemilik mengetahui kemungkinan penyakit yang menyerang kucingnya sehingga dapat langsung melakukan penanganan. Dengan adanya sebuah sistem yang dapat mengolah data terkait masalah ini, secara khusus dapat membantu dalam menginformasikan hal-hal terkait diagnosa penyakit pada kucing.

Oleh karena itu, memilih metode atau teknik yang baik dan tepat selama pengerjaan untuk penyelesaian masalah adalah hal yang penting. Referensi yang baik akan membantu mengurangi beban kerja pengembang pada langkah selanjutnya. Maka pada penelitian ini dapat memanfaatkan salah satu metode yang merupakan implementasi dari data mining yaitu metode *Modified K-Nearest Neighbor* (MKNN). *Modified K-Nearest Neighbor* (MKNN) merupakan metode yang dapat digunakan untuk diagnosa penyakit pada kucing. *Modified K-Nearest Neighbor* (M-KNN) adalah pengembangan dari metode KNN yang dirancang untuk mengatasi kelemahan dari jarak data dengan *weight* pada *outlier* KNN. (Prasetyo, 2012) Dengan metode ini, sistem dapat melakukan perhitungan dan melakukan diagnosa terhadap penyakit tersebut.

Salah satu contoh implementasi dari metode *Modified K-Nearest Neighbor* (MKNN) dari penelitian "Diagnosis penyakit tanaman tomat menggunakan metode *Modified K-Nearest*

Neighbor" oleh Prasiwi Meilida Basuki dkk. Penelitian ini membahas tentang bagaimana menerapkan metode *Modified K-Nearest Neighbor* dalam mendiagnosa penyakit pada tanaman tomat. Pada penelitian ini menggunakan 6 jenis penyakit dengan 15 gejala penyakit tanaman tomat yang terlihat pada batang, daun dan buah. Pada pengujian dengan nilai yang berbeda diperoleh nilai rata-rata akurasi yang berbeda pula pada masing-masing data latih. Data latih yang digunakan dalam pengujian adalah data latih 25%, data latih 50% dan data latih 75%. Nilai akurasi terbesar ditunjukkan pada nilai $k=1$ dengan rata-rata akurasi sebesar 98,92%. Sedangkan untuk akurasi terkecil pada nilai $k=10$ dengan rata-rata akurasi 89,04%. Dengan peningkatan jumlah data latih berpengaruh terhadap peningkatan nilai akurasi. (Meilida, 2015).

Berdasarkan permasalahan yang telah diuraikan serta melihat dari penelitian yang dilakukan sebelumnya maka judul yang diusulkan dalam penelitian ini adalah "Implementasi Algoritma *Modified K-Nearest Neighbor* (MKNN) untuk Diagnosa Penyakit pada Kucing". Sehingga dengan adanya penelitian ini diharapkan dapat membantu dalam menyelesaikan permasalahan diagnosa penyakit pada kucing

1.2 Rumusan Masalah

Berdasarkan latar belakang yang telah dibahas sebelumnya, maka rumusan masalah yang dapat dikaji adalah sebagai berikut :

1. Bagaimana merancang dan mengimplementasi algoritma *Modified K-Nearest Neighbor* (M-KNN) untuk klasifikasi gejala diagnosa penyakit pada kucing?
2. Bagaimana tingkat akurasi, *precision* dan *recall* menggunakan algoritma *Modified K-Nearest Neighbor* (M-KNN) untuk diagnosa penyakit pada kucing?

1.3 Tujuan

Tujuan dari penelitian "Implementasi Algoritma *Modified K-Nearest Neighbor* (MKNN) untuk Diagnosa Penyakit pada Kucing" sebagai berikut:

1. Membuat dan merancang implementasi dalam mendiagnosa penyakit pada kucing dengan menggunakan metode *Modified K-Nearest Neighbor* (MKNN).
2. Menguji tingkat akurasi, *precision*, dan *recall* algoritma *Modified K-Nearest Neighbor* (M-KNN) untuk klasifikasi gejala terhadap diagnosa penyakit pada kucing.

1.4 Batasan Masalah

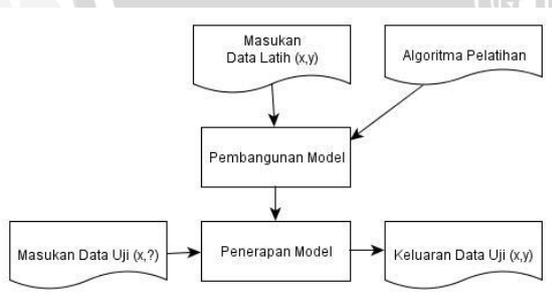
Adapun batasan masalah dalam penelitian ini sebagai berikut:

1. Pengembangan implementasi untuk mendiagnosa penyakit pada kucing menggunakan metode *Modified K-Nearest Neighbor* (MKNN).
2. Identifikasi berdasarkan gejala yang terlihat pada kucing.
3. Data yang digunakan sebanyak 220 data kucing yang terkena penyakit.
4. Sistem nantinya akan menghasilkan keluaran berupa jenis penyakit dimana pada sistem ini hanya akan mengidentifikasi 18 gejala dan 8 jenis penyakit yaitu *Scabies, Gastritis, Helminthiasis, Rhinitis, Dermatophytosis, dermatitis, enteritis* dan *Otitis*.
5. Sistem dibangun dengan menggunakan bahasa pemrograman PHP.

2. Landasan Kepustakaan

2.1 Klasifikasi

Memberikan nilai pada objek data agar data tersebut masuk ke kelas tertentu dari kelas yang ada disebut dengan klasifikasi. Terdapat 2 pekerjaan utama yang dilakukan dalam klasifikasi yaitu membangun model yang digunakan sebagai prototipe untuk disimpan sebagai memori dan menggunakan model sebagai pengenalan/klasifikasi/prediksi untuk mengetahui letak kelas dari objek tersebut.



Gambar 2.2 Proses pekerjaan klasifikasi (Presetyo, 2012)

2.2 K-Nearest Neighbor (K-NN)

Proses utama yang dilakukan adalah dengan memberikan satu set N data (data training) yang telah diketahui kelasnya. Kemudian mengklasifikasikan satu set N data (data testing) berdasarkan keterdekatannya pada setiap data training dengan mengacu nilai k yang telah ditentukan sebelumnya. (Nitin, 2010)

Kelas yang muncul pada pengelompokan tersebut akan dibandingkan jumlahnya yang mana kelas yang memiliki jumlah yang paling besar

merupakan hasil klasifikasi (Seyyid, 2013). Secara sederhana dari algoritma KNN adalah implementasi dari Brute Force yang terdiri dari 3 tahap. Tahap pertama adalah untuk menghitung semua jarak dari setiap titik pada data training dengan data uji. Tahap kedua adalah untuk memilah jarak tersebut berdasarkan nilai k yang telah ditentukan. Dan tahap ketiga adalah menentukan kelas hasil klasifikasi berdasarkan jumlah kelas yang telah dikelompokkan berdasarkan nilai k tersebut. (Otair, 2013)

2.2.1 Penentuan Jarak

Dalam matematika, jarak adalah sebuah bentuk yang menyatakan jarak panjang antara dua titik. Jarak tersebut yang dapat kita gunakan untuk mengelompokkan individu yang sama dan memisahkan yang tidak serupa. Cara dalam menentukan jarak dapat menggunakan *eclidean distance*. (Seyyid, 2013)

$$d(X_i, X_j) = \sqrt{\sum_{r=1}^n (X_{ir} - X_{jr})^2}$$

Untuk data kontinu dapat melakukan proses normalisasi untuk pemerataan data agar rentangnya tidak terlalu besar. Perhitungan min-max dilakukan untuk mengubah atribut A dengan nilai v menjadi v' dalam range [0,1]. Perhitungan dapat dilihat pada persamaan 2-2. (Han & Kamber, 2006) :

$$v' = \frac{v - \min_A}{\max_A - \min_A} (new_{\max_A} - new_{\min_A}) + new_{\min_A}$$

Keterangan:

- v' : Nilai normalisasi
- v : Nilai data yang dinormalisasi
- min_A : Nilai terendah data pada atribut A
- max_A : Nilai tertinggi data pada atribut A
- new_{minA} : Nilai minimum terbaru
- new_{maxA} : Nilai maksimum terbaru

2.3 Modified K-Nearest Neighbor (MKNN)

Metode *Modified K-Nearest Neighbor* (MKNN) merupakan sebuah metode hasil dari modifikasi metode *K-Nearest Neighbor* (KNN). Dalam metode ini, setiap sampel data latih harus divalidasi pada langkah pertama. Validasi pada setiap titik dihitung sesuai dengan nilai k tetangganya. Modifikasi metode KNN dilakukan untuk meningkatkan kinerja metode tersebut dan meningkatkan akurasi. Pada algoritma KNN belum ada pembobotan untuk setiap dataset. Metode *Modified K-Nearest Neighbor* (MKNN) dilakukan



dengan penambahan perhitungan nilai validasi dan perhitungan *weight voting*. (Hamid, 2010)

Validasi pada data bergantung pada nilai k terdekat dimana tujuan perhitungan ini adalah untuk menghitung jumlah titik dengan label sama pada data latih. (Parvin, 2010)

$$Validity(x) = \frac{1}{H} \sum_{i=1}^H S(lbl(x), lbl(N_i(x)))$$

Keterangan:

- H : Jumlah titik terdekat
- Lb(x) : Kelas x
- $N_i(x)$: Label kelas titik terdekat x
- S : Menghitung kesamaan antara titik a dan data ke-b k terdekat.

Pada algoritma *Modified K-Nearest Neighbor* (M-KNN), *Weight voting* dihitung dengan persamaan berikut ini: (Parvin, 2010)

$$W(i) = Validity(i) \times \frac{1}{d_e + 0.5}$$

- W(i) : Weigh Voting
- Validasi(x) : Nilai Validasi
- d_e : Jarak Euclidean

2.4 Akurasi, Precision dan Recall

Untuk mengetahui nilai akurasi, *precision* dan *recall* pada pengujian yang akan dilakukan maka dapat menggunakan matrik konfusi pada tabel 2.1. (Prasetyo, 2012)

Tabel 2.1 Matriks konfusi klasifikasi 2 kelas

fij		Kelas hasil prediksi	
		Kelas=1	Kelas=0
Kelas asli (i)	Kelas=1	f_{11}	f_{10}
	Kelas=0	f_{01}	f_{00}

$$Akurasi = \frac{f_{11} + f_{00}}{f_{11} + f_{10} + f_{01} + f_{00}}$$

Pada sel *fij* merupakan jumlah record atau data dari kelas i yang hasil prediksi masuk ke kelas j. Seperti, pada sel f_{11} merupakan jumlah data dalam kelas 1 yang dengan benar dipetakan ke kelas 1 sedangkan pada sel f_{10} merupakan data dalam kelas 1 yang dipetakan salah ke kelas 0. (Prasetyo, 2012)

$$Precision = \frac{f_{11}}{f_{11} + f_{01}}$$

$$Recall = \frac{f_{11}}{f_{11} + f_{10}}$$

3. METODOLOGI

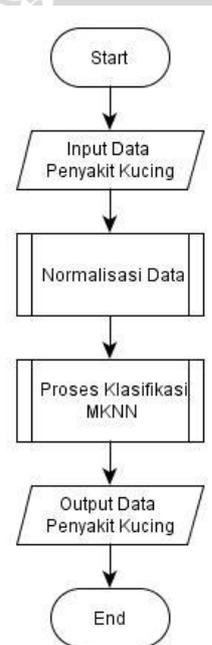
Metodologi penelitian memuat tentang langkah-langkah dalam membangun implementasi algoritma *Modified K-Nearest Neighbor* (MKNN) untuk klasifikasi diagnosa penyakit pada kucing.



Gambar 3.1 Tahapan metodologi penelitian

4. PERANCANGAN

Bab ini menjelaskan tentang analisis kebutuhan dan perancangan untuk membangun sebuah sistem dalam memberikan informasi kepada pengguna mengenai diagnosa penyakit kucing. Klasifikasi pada sistem ini menggunakan metode *Modified K-Nearest Neighbor* (MKNN). Kriteria pada preproses data ini menggunakan 18 gejala dari 8 penyakit.



Gambar 4.1 Diagram alir sistem



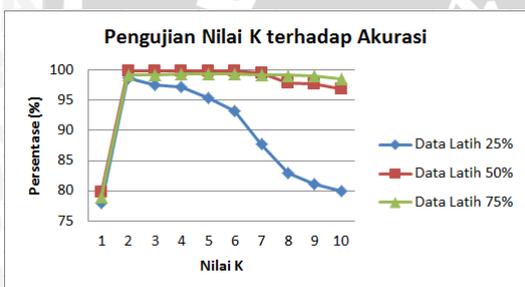
5. PENGUJIAN DAN ANALISIS

Pengujian yang dilakukan pada penelitian ini adalah pengujian pengaruh nilai k, pengujian pengaruh sebaran data dan pengujian pengaruh jumlah data latih. Pada pengujian ini dilakukan 5 kali uji coba dan diambil rata-rata dari uji coba tersebut.

5.1 Pengujian Pengaruh Nilai k

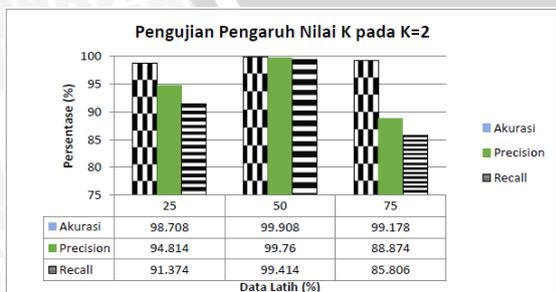
Pada pengujian ini dilakukan sebanyak 3 kali uji coba yaitu pada data latih 25%, data latih 50%, dan data latih 75% dengan cara memberikan nilai k dari 1 hingga 10. Pada data latih 50% menggunakan 110 data latih dan 110 data uji. Sedangkan pada data latih 75% menggunakan 165 data latih dan 55 data uji.

Pada pengujian ini akurasi pada grafik akan cenderung menurun dari k=2 hingga k=10 karena jika nilai k semakin besar maka semakin banyak pula ketetangaan terdekat yang akan digunakan dalam proses klasifikasi. Dari hal tersebut semakin banyak ketetangaan yang seharusnya tidak dipertimbangkan menjadi pertimbangan dan masuk kedalam



Gambar 5.1 Grafik hasil pengujian perubahan nilai k terhadap akurasi

Pada proses perhitungan, nilai akurasi tertinggi terletak pada nilai k=2 pada data latih 50% dengan akurasi 99,90 %, precision 99,76% dan recall 99,414%. Gambar 5.2 menunjukkan pengaruh nilai k pada k=2.



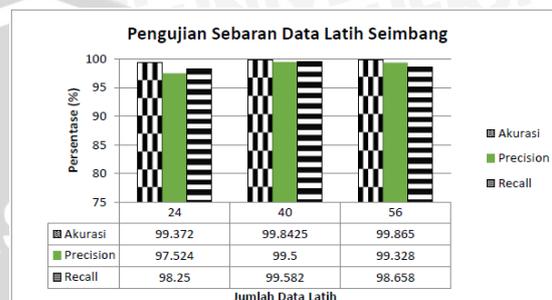
Gambar 5.2 Grafik hasil pengujian perubahan nilai k pada k=2

5.2 Pengujian terhadap Sebaran Data.

Pengujian ini terdiri dari pengujian jenis data seimbang dan tidak seimbang dengan menggunakan jumlah data uji dan data latih sama

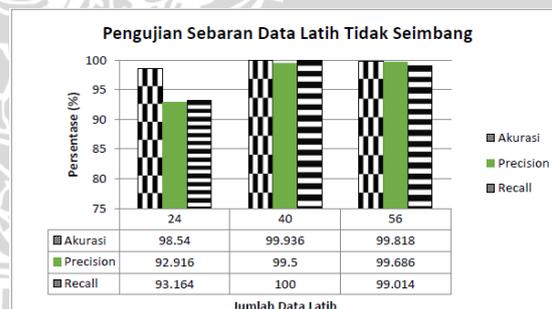
serta menggunakan nilai k=2 yang didapatkan dari pengujian pertama.

Pada jenis data latih yang seimbang terlihat bahwa grafik cenderung meningkat seiring dengan bertambahnya jumlah data latih yang digunakan. Jenis data latih seimbang merupakan jumlah penyakit (kelas) yang digunakan berjumlah sama pada masing-masing data latih yang digunakan terlihat pada gambar 5.3.



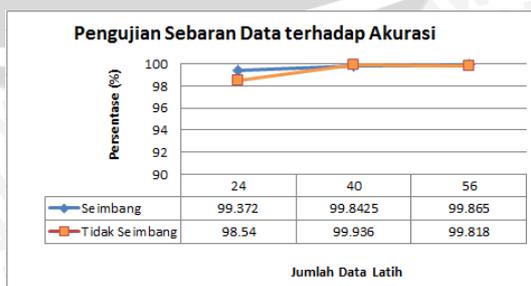
Gambar 5.3 Grafik Pengujian Sebaran Data Seimbang

Sedangkan pada jenis data latih tidak seimbang mengalami kenaikan namun kurang stabil. Jenis data latih tidak seimbang merupakan jumlah kelas yang digunakan dalam satu kali uji coba berbeda ditunjukkan pada gambar 5.4.



Gambar 5.4 Grafik Pengujian Sebaran Data Tidak Seimbang

Hal ini dikarenakan pada jenis data latih tidak seimbang ini, terdapat dominasi dari penyakit (kelas) tertentu yang dapat menyebabkan hasil yang tidak akurat dalam proses klasifikasi. Proses klasifikasi dipengaruhi oleh penyakit (kelas) mendominasi.



Gambar 5.5 Grafik Pengujian Sebaran Data terhadap Akurasi

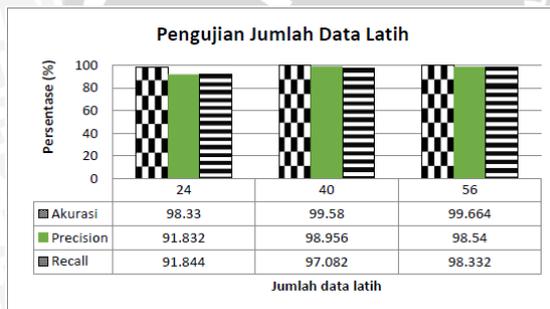
5.3 Pengujian terhadap Pengaruh Jumlah Data Latih.

Pengujian ini dilakukan dengan 3 uji coba yaitu dengan 24, 40, dan 56 data latih dengan data uji berjumlah 30 ditunjukkan pada tabel 5.1. Pengujian ini dilakukan untuk melihat pengaruh peningkatan data latih.

Tabel 5.1 Hasil pengujian ketiga

Data Latih	Data Uji	Akurasi (%)	Precision (%)	Recall (%)
24	30	98.33	91.844	91.832
40		99.58	97.082	98.956
56		99.664	98.332	98.54

Dari pengujian yang dilakukan terlihat bahwa akurasi semakin meningkat mencapai angka 99,66 % dengan presicion 98,33 % dan recall 98,54% pada 56 data latih ditunjukkan oleh gambar grafik 5.6. Hal ini dikarenakan semakin banyak data latih yang digunakan pada proses perhitungan maka proses *learning* pada metode yang digunakan semakin baik.



Gambar 5.6 Grafik Pengujian Jumlah Data Latih

6. PENUTUP

6.1 Kesimpulan

Berdasarkan pada pengujian yang telah dilakukan terhadap diagnosa penyakit kucing menggunakan metode Modified K-Nearest Neighbor (MKNN) dapat diambil kesimpulan sebagai berikut:

1. Algoritma *Modified K-Nearest Neighbor* (MKNN) dapat diimplementasikan untuk diagnosa penyakit pada kucing dengan menggunakan 18 parameter yaitu gejala penyakit yang menyerang kucing dengan 8 jenis penyakit.
2. Dari hasil pengujian yang telah dilakukan menghasilkan kesimpulan sebagai berikut:
 - a. Berdasarkan hasil pengujian pengaruh nilai k didapatkan hasil akurasi maksimum 99,90 %, akurasi minimum 79,76%, *precision* 99,76% dan *recall* 99,41% pada data latih 50% yang terdiri dari 110 data latih dan 110 data uji. Dari hasil pengujian menunjukkan bahwa perubahan nilai k berpengaruh pada akurasi. Nilai k terbaik

dipengaruhi oleh data yang digunakan. Dengan adanya penambahan nilai k maka grafik akurasi akan cenderung menurun karena semakin banyak ketetanggaan terdekat yang akan digunakan dalam proses klasifikasi, semakin banyak pula ketetanggaan yang seharusnya tidak dipertimbangkan menjadi pertimbangan dan masuk kedalam proses perhitungan.

- b. Berdasarkan hasil pengujian sebaran data, akurasi yang dihasilkan oleh jenis data latih seimbang mengalami peningkatan sedangkan pada jenis data latih tidak seimbang peningkatan terjadi kurang stabil. Hal ini dikarenakan pada data latih tidak seimbang proses klasifikasi yang terjadi akan cenderung pada kelas yang mendominasi.
- c. Berdasarkan hasil pengujian jumlah data latih, akurasi mengalami peningkatan dikarenakan proses *learning* pada metode yang digunakan semakin baik.

6.2 Saran

Saran yang dapat diberikan dari penelitian "Implementasi algoritma Modified K-Nearest neighbor untuk klasifikasi diagnosa penyakit kucing" untuk penelitian selanjutnya sebagai berikut:

1. Untuk penelitian mendatang, sebaiknya melakukan penambahan terhadap parameter maupun jumlah data untuk mengoptimalkan hasil dari proses klasifikasi. Penambahan parameter dapat berupa jenis penyakit maupun gejala penyakit.
2. Peneliti selanjutnya dapat mengembangkan sistem dengan cara mengkombinasikan dengan metode lain agar mendapatkan hasil yang lebih optimum dan efektif.

DAFTAR PUSTAKA

- Hamid Parvin, 2010. *A Modification on K-Nearest Neighbor Classifier*, Global Journal of Computer Science and Technology.
- Han, J. & Kamber, M., 2006. *Data Mining : Concepts and Techniques 2nd ed.* San Fransisco: Morgan Kauffman Publisher.
- Holbrook B., Morris, 2008. *Pets and people: Companions in commerce*, Journal of Business Research 61 546–552.
- Meilida Prasiwi, Marji, & Nurul, 2015. *Diagnosis Penyakit Tanaman Tomat Menggunakan Algoritma Modified K-Nearest Neighbor (MKNN)*.S1. Universitas Brawijaya.

Nitin Bhatia, Vandana, 2010. *Survey of Nearest Neighbor Condensing Techniques*. International Journal of Computer Science and Information Security Vol.8 No.2.

Otair Muhammed, 2013. *Approximate k-Nearest Neighbor Based Spatial Clustering using K-D Tree*. International Journal of Database Management Systems Vol.5 No.1.

Parvin, H., Alizadeh, H. & Minati, B., 2010. *A Modification on K-Nearest Neighbor Classifier*. *Global Journal of Computer Science and Technology*, Vol.10 (Issue 14).

Prasetyo Eko, 2012. *Data Mining : Konsep dan Aplikasi Menggunakan MATLAB*. Yogyakarta: Andi.

Setianingrum, 2012. *Manfaat Memelihara Hewan pada Penderita Penyakit Kronis*. Fakultas Psikologi. Universitas Muhammadiyah Malang, Malang.

Seyyid Ahmed Medjahed, 2013. *Breast Cancer Diagnosis by Using k-Nearest Neighbor with Different Distances and Classification Rule*, International Journal of Computer application vol 62 No.1.

