

METODE ENSEMBLE CLASSIFIER UNTUK MENDETEKSI JENIS ATTENTION DEFICIT HYPERACTIVITY DISORDER (ADHD) PADA ANAK USIA DINI

Indriati¹, Ari Kusyanti²

^{1,2}Jurusan Teknik Informatika Fakultas Ilmu Komputer Universitas Brawijaya
Email: ¹indriati.tif@ub.ac.id, ²ari.kusyanti@ub.ac.id

(Naskah masuk: 29 November 2018, diterima untuk diterbitkan: 13 Februari 2019)

Abstrak

Pada awal masa perkembangan, beberapa anak mengalami hambatan diantaranya sulit untuk diam, sulit untuk berkonsentrasi dan mengontrol perilakunya, apabila anak mengalami gangguan pemusatan perhatian dan sulit mengontrol perilaku yang sesuai, dapat disebut dengan ADHD (*Attention Deficit Hyperactive Disorder*). Ini merupakan masalah yang serius dikarenakan anak penyandang ADHD mengalami masalah perilaku sosial, emosional dan mengalami kesulitan belajar sekolah sehingga akan mempengaruhi perkembangan pada masa dewasa anak penyandang ADHD. Oleh karena itu perlu diketahui gejala ADHD sejak dini, agar dapat dilakukan suatu penanganan dengan cepat dan tepat. Penelitian ini menghasilkan aplikasi yang digunakan untuk mendeteksi jenis ADHD berdasarkan gejala-gejala yang di masukkan oleh pengguna sehingga akan tampil hasil klasifikasi jenis ADHD nya secara otomatis. Aplikasi ini menggunakan metode *Ensemble Classifier* yaitu metode yang menggabungkan beberapa *classifier* agar dapat meningkatkan akurasi yang dihasilkan. Pada tahap klasifikasi setiap data akan dihitung menggunakan *K-Nearest Neighbour* (KNN), *Fuzzy K-Nearest Neighbour* (FKNN) dan *Neighbour Weighted K-Nearest Neighbour* (NWKNN). Hasil perhitungan ketiga *classifier* tersebut akan diproses kembali dengan metode *Ensemble Classifier* dengan menggunakan *majority voting* untuk penentuan klasnya. Hasil akurasi tertinggi dari metode *ensemble classifier* yaitu 95% dengan nilai k optimal yaitu k=10. Akan tetapi semakin besar nilai k yaitu diatas k=20 maka nilai akurasi untuk masing-masing algoritme akan semakin turun. Hal ini dikarenakan semua algoritme penentuan klasifikasinya berdasarkan jumlah ketetanggaannya. Maka semakin banyak jumlah tetangga yang diperhitungkan maka kemungkinan salah klasifikasinya semakin besar.

Kata kunci: ADHD, KNN, FKNN, NWKNN, *Ensemble Classifier*

ENSEMBLE CLASSIFIER METHOD FOR DETECTING TYPE OF ATTENTION DEFICIT HYPERACTIVITY DISORDER (ADHD) IN EARLY AGE

Abstract

At the beginning of the development stage, some children experience difficulty to calm, to concentrate and to control their behavior. These symptoms are known as ADHD (*Attention Deficit Hyperactive Disorder*). This research develops an application that is used to define ADHD based on symptoms that that is entered by the user so that it will show its ADHD type automatically. This application uses the *Ensemble Classifier* method, in which a method that allows some classifier in order to increase the resulting value. At the classification stage each data will be calculated using *K-Nearest Neighbor* (KNN), *Fuzzy K-Nearest Neighbor* (FKNN) and *Neighbor Weighted K-Nearest Neighbor* (NWKNN). The results of the three classifier calculations will return using the *Ensemble Classifier* method using the *majority voting* for class determination. Acceptance results from the ensemble classifier method is 95% with the optimal k value k = 10. However, when the k value, i.e k >=20 then the value for each algorithm will decrease. This is due to the calculation of all the classification algorithm based on the number of its neighbors. Therefore, the more neighbours that are calculated then the possibility of misclassification is greater.

Keywords: ADHD, KNN, FKNN, NWKNN, *Ensemble Classifier*

1. PENDAHULUAN

Kanak-kanak awal merupakan masa dimana pertumbuhan dan perkembangan anak sangat pesat. Masa ini berlangsung dari usia 2 hingga 6

tahun, sering disebut juga *golden age* (masa keemasan). Pada masa keemasan anak ini perlu adanya stimulus atau rangsangan awal untuk memunculkan perkembangan yang sesuai dengan tahapan usianya. Tanggapan terhadap rangsangan

akan dipergunakan sebagai pemantau apabila terjadi hambatan atau permasalahan dalam proses perkembangan pada anak. Hurlock menjelaskan bahwa anak sangat berpotensi dalam mengalami hambatan perkembangan selama masa keemasan tersebut, akan tetapi tidak menutup kemungkinan akan terjadi hambatan perkembangan di luar usia 2-6 tahun (Hurlock, 1980).

Hambatan perkembangan ini muncul pada masa perkembangan usia anak kurang dari 7 tahun. Pada usia tersebut, ADHD sudah dapat terdeteksi karena ketidaksesuaian perkembangan yang muncul dengan perkembangan anak seusianya (Nasution, 2015). Kasus ADHD sendiri dua kali lipat ditemukan pada anak laki-laki dibandingkan anak perempuan. Di Indonesia sendiri kasus ADHD masih belum terlalu banyak dikenali. Data menyebutkan bahwa sekitar 2-11 % anak-anak pada usia sekolah yang ada di dunia mengalami ADHD (Dina dkk, 2009). Oleh karena itu perlu diketahui gejala ADHD sejak dini, agar dapat dilakukan suatu penanganan dengan cepat dan tepat.

Seringkali orang tua menganggap gejala yang muncul pada hambatan perkembangan tersebut tergolong wajar karena anak masih dalam masa perkembangan. Beberapa orang tua juga belum terlalu memahami gejala yang dialami anak penyandang ADHD. Ini merupakan masalah yang serius dikarenakan anak penyandang ADHD mengalami masalah perilaku sosial, emosional dan mengalami kesulitan belajar sekolah. Hal tersebut akan mempengaruhi perkembangan pada masa dewasa anak penyandang ADHD (Dafison dkk, 2006).

Permasalahan tersebut dapat diatasi dengan membuat sistem sebagai alat bantu yang digunakan untuk mendeteksi jenis ADHD berdasarkan gejala-gejala yang di masukkan oleh pengguna. Sistem tersebut akan membutuhkan pengetahuan mengenai kecerdasan buatan (*artificial intelligence*). Pada proses implementasi kecerdasan buatan bisa berupa Data Mining dan Pengenalan Pola yang meliputi klasifikasi dan clustering, Sistem Pakar, maupun Sistem Pendukung Keputusan. Dari beberapa proses implementasi kecerdasan buatan tersebut, klasifikasi merupakan proses yang dapat membantumendeteksi jenis ADHD. Sistem akan melakukan klasifikasi otomatis berdasarkan gejala-gejala yang di masukkan oleh pengguna.

Selama ini pengklasifikasian banyak berfokus pada *single classifier*, namun dengan menggunakan metode *Ensemble* yaitu metode yang menggabungkan beberapa *classifier* dapat meningkatkan akurasi yang dihasilkan (Nadia, 2014). Karena setiap *classifier* memiliki kelebihan yang berbeda-beda, dan jika beberapa *classifier* tersebut dikombinasi tentu hasil yang dicapai akan lebih baik. Keunggulan menggunakan metode *Ensemble* adalah hasil klasifikasi lebih akurat karena metode ini akan membandingkan hasil antara *single classifier* yang digunakan.

Berdasarkan penelitian-penelitian terdahulu maka penelitian ini akan membahas mengenai bagaimana menggunakan metode *ensemble classifier* untuk mendeteksi jenis ADHD. Selain itu akan dibahas juga mengenai hasil pengujian metode *ensemble classifier* dengan menguji berdasarkan nilai k dan variasi perubahan jumlah data latih dan data uji untuk mengetahui parameter optimal yang harus digunakan untuk menghasilkan nilai akurasi yang tinggi. Diharapkan dengan menggunakan metode *ensemble classifier* menghasilkan akurasi yang lebih baik daripada *single classifier*

2. ATTENTION DEFICIT HYPERACTIVE DISORDER (ADHD)

ADHD merupakan hambatan pada proses perkembangan, yang ditandai dengan kesulitan untuk memusatkan perhatian, pembahasan ini tercantum dalam DSM IV. Hambatan perkembangan ini mempunyai karakteristik kurangnya pemusatan perhatian, impulsif dan hiperaktif sehingga mempengaruhi aktifitas sehari-hari. Akibat kurang mampunya dalam pemusatan perhatian (*inattention*), akan mempengaruhi proses penerimaan informasi sehingga menjadi kurang optimal. Hal tersebut menimbulkan permasalahan dalam perkembangan usia sekolah. Individu yang mengalami ADHD pada usia kanak-kanak, dapat berlangsung hingga masa dewasa (Neale, J.M, 2014).

Menurut DSM-IV jenis ADHD sendiri adalah sebagai berikut (Neale, 2004):

INATTENTION

- Sering gagal fokus dalam memberi perhatian atau membuat kesalahan yang tidak terkontrol dalam sekolah, bekerja dan aktifitas sehari-hari
- Sering mengalami kesulitan menjaga perhatian atau konsentrasi dalam menerima tugas atau aktifitas bermain.
- Sering kelihatan tidak mendengarkan ketika diajak berbicara secara langsung
- Sering kali menghindar, tidak senang atau enggan mengerjakan tugas yang membutuhkan usaha (seperti pekerjaan sekolah atau pekerjaan rumah)
- Sering kehilangan barang-barang yang dibutuhkan untuk tugas atau kegiatan (permainan, tugas sekolah, pensil, buku dan alat sekolah lainnya)
- Sering kali mengalihkan perhatian dari rangsangan dari luar yang tidak berkaitan
- Sering kali melupakan tugas atau kegiatan sehari-hari

HIPERAKTIFITAS

- Sering gelisah yang tampak pada tangan dan kaki serta sering menggeliat dalam tempat duduk.
- Sering tidak dapat berdiam lama di tempat duduk dalam kelas atau situasi lain yang mengharuskan tetap duduk.
- Sering berlari-larian atau memanjat secara berlebihan dalam situasi yang tidak seharusnya

- (pada dewasa atau remaja biasanya terbatas dalam keadaan perasaan tertentu atau kelelahan)
- d. Sering kehilangan ketenangan dalam mengisi waktu luang dan kesulitan bermain
 - e. Sering kali berperilaku seperti mengendarai motor
 - f. Seringnya berbicara secara berlebihan

IMPULSIF

- a. Sering berbicara tanpa berpikir dan menjawab pertanyaan sebelum pertanyaannya selesai.
- b. Sering tidak sabar menunggu giliran atau antrian
- c. Sering berbicara untuk menyela atau memaksakan kemauan diri terhadap orang lain (misalnya dalam percakapan atau permainan)

3. ENSEMBLE CLASSIFIER

Klasifikasi merupakan metode untuk menemukan suatu fungsi maupun model yang menjelaskan dan membedakan konsep suatu kelas pada data, yang tujuannya untuk memprediksi kelas dari suatu objek yang labelnya belum diketahui. Secara umum, proses klasifikasi dibagi menjadi dua fase, yang pertama adalah, *learning* pada fase ini, data yang telah diketahui kelas pada datanya akan di umpangkan untuk membentuk suatu model perkiraan. Kedua, *test* pada fase ini model yang telah terbentuk, diuji untuk mengetahui akurasi dari data tersebut. Pada klasifikasi ada beberapa jenis metode, yaitu *support vector machine* (SVM), *k-nearest neighbor*, *decision tree*, *Bayesian*, *fuzzy* dan *neural network* (Pramudiono, 2003).

Metode *Ensemble* adalah penggabungan beberapa *single classifier* untuk membentuk sebuah classifier baru sehingga didapatkan hasil yang lebih akurat (Verbaeten, 2003). Metode *Ensemble* telah banyak digunakan dalam beberapa penelitian karena telah terbukti didapatkan hasil akurasi yang lebih baik (Da Silva, 2014). Dan dapat mengurangi varians dalam kesalahan perhitungan oleh pengklasifikasian tunggal (Breiman, 1996). Metode *Ensemble* menggabungkan beberapa *single classifier* dengan tujuan keunggulan dari masing-masing classifier akan dapat digabung sehingga performa yang didapat dalam memecahkan masalah menjadi lebih baik. Langkah pengerjaan metode ini adalah dengan menghitung hasil dari masing-masing classifier kemudian akan dibobotkan kembali menggunakan pembobotan *Ensemble* yaitu *majority voting*, *average of probability* dan *weighted voting*.

3.1. K-Nearest Neighbor (KNN)

KNN (K-Nearest Neighbor) adalah salah satu metode klasifikasi yang menggunakan kedekatan ketetanggaan dengan data latih untuk menentukan kelas data uji (Adeniyi et al, 2016). Menurut D.A Adeniyi et al (2016) tahapan proses pada algoritme KNN diawali dengan menentukan nilai parameter K kemudian menghitung nilai kedekatan ketetanggaan antara data uji dengan data latih dengan

menggunakan persamaan jarak *Euclidean Distance* atau persamaan kedekatan dengan *Cosine Similarity* Perhitungan jarak dengan *Euclidean Distance* dapat diformulasikan pada Persamaan 1.

$$d(x_1, x_2) = \sqrt{\sum_{i=1}^n (x_{2i} - x_{1i})^2} \tag{1}$$

Keterangan :

X_1 : nilai data dari tiap data latih

X_2 : nilai data dari tiap data uji

n : banyaknya data

i : data ke-i

Sedangkan untuk perhitungan *Cosine Similarity* dapat diformulasikan pada Persamaan 2

$$CosSim(q, d_j) = \frac{\vec{d}_j \cdot \vec{q}}{|\vec{d}_j| \cdot |\vec{q}|} = \frac{\sum_{i=1}^m (w_{ij} \cdot w_{iq})}{\sqrt{\sum_{i=1}^m w_{ij}^2} \cdot \sqrt{\sum_{i=1}^m w_{iq}^2}} \tag{2}$$

Keterangan:

$CosSim(q, d_j)$: nilai Cosine Similarity, yaitu

nilai similitas data uji q dengan data latih j

q : nilai data uji

d_j : nilai data latih

$\vec{d}_j \cdot \vec{q}$: hasil total perkalian vektor

antara data latih dengan data uji

$|\vec{d}_j| \cdot |\vec{q}|$: hasil total perkalian vektor

antara nilai normalisasi data latih dengan data uji

w_{ij} : bobot nilai i pada data latih j

w_{iq} : bobot nilai i pada data uji

m : banyaknya jumlah nilai

Setelah didapatkan hasil kedekatan ketetanggaan, kemudian dilakukan pengurutan hasil *Euclidean Distance* dari hasil terkecil ke terbesar atau mengurutkan hasil *CosSim* dari terbesar ke terkecil. Ketika telah didapatkan hasil K tetangga terdekat dari hasil penurutan tersebut maka dihitung nilai skornya berdasarkan nilai yang ada, perhitungan skor tersebut dapat di formulasikan kedalam Persamaan 3 dan 4

$$Score(X, C_i) = \sum_{djKNN(X)} \left(\left(\sqrt{\sum_{i=1}^n (x_{2i} - x_{1i})^2} \right) * \delta(d_j, C_i) \right) \tag{3}$$

atau

$$Score(X, C_i) = \sum_{djKNN(X)} \left((Sim(q, d_j) * \delta(d_j, C_i)) \right) \tag{4}$$

Dimana :

$djKNN(x)$: data latih d_j pada kumpulan

tetangga

terdekat dari data uji X

$\sqrt{\sum_{i=1}^n (x_{2i} - x_{1i})^2}$: jarak antara data uji dan data latih

- $\delta(d_j, C_i)$: akan bernilai 1 jika $d_j \in C_i$ dan bernilai 0 jika $d_j \notin C_i$
- C_i : jenis atau kelas i
- $Sim(q, d_j)$: nilai *CosSim* antara data uji dan data latih.

3.2. Fuzzy K-Nearest Neighbor (FKNN)

Fuzzy K-Nearest Neighbor (FK-NN) merupakan metode yang menggabungkan antara dua metode klasifikasi yaitu Logika Fuzzy dan *K-Nearest Neighbor*. Algoritme *Fuzzy K-Nearest Neighbor* memiliki 2 kelebihan jika dibandingkan dengan metode *K-Nearest Neighbor*, yaitu dapat membantu proses klasifikasi jika ada sifat ambigu dari data tetangga misalnya dengan nilai jarak atau kedekatan yang sama. Pada FK-NN, setiap data akan memiliki nilai derajat keanggotaan pada masing-masing kelas sehingga akan lebih memberikan kekuatan/kepercayaan suatu data berada pada suatu kelas. Sifat ambigu dari tetangga akan memiliki pengaruh penting dalam penentuan kelas agar tidak salah klasifikasi (Jowik, 2013).

Tahapan-tahapan proses yang harus dilakukan untuk menerapkan algoritme *Fuzzy K-Nearest Neighbor* (FK-NN) adalah sebagai berikut :

1. Menghitung jarak kedekatan antara data uji dengan data latih dengan menggunakan Persamaan 5

$$d_{(x_i, x_j)} = (\sum_{l=1}^N |(x_{il} - x_{jl})|^2)^{\frac{1}{2}} \tag{5}$$

Keterangan:

- $d_{(x_i, x_j)}$: jarak antar titik pada data uji dan data latih
- N : jumlah fitur data
- x_{il} : data uji ke- l
- x_{jl} : data latih ke- l

2. Menghitung nilai keanggotaan untuk setiap data latih menggunakan Persamaan 6 untuk setiap i ,

$$1 \leq i \leq c$$

$$u_{(x, c_i)} = \frac{\sum_{k=1}^K u_{(x_k, c_i)} * d_{(x, x_k)}^{\frac{-2}{(m-1)}}}{\sum_{k=1}^K d_{(x, x_k)}^{\frac{-2}{(m-1)}}} \tag{6}$$

Keterangan:

- $u_{(x, c_i)}$: nilai keanggotaan dari data x terhadap kelas c_i
- k : jumlah tetangga terdekat
- $u_{(x_k, c_i)}$: nilai keanggotaan dari data tetangga dalam k tetangga pada kelas c_i
- $d_{(x, x_k)}$: jarak dari data x ke data x_k dalam k tetangga terdekat
- m : bobot pangkat, $m > 1$

4. Mengambil nilai terbesar dari nilai keanggotaan untuk semua $1 \leq i \leq c$
5. Memberi label kelas c ke data uji- x .

3.3. Neighbor Weighted K-Nearest Neighbor (NWKNN)

NWKNN merupakan salah satu algoritme yang hampir mirip dengan algoritme KNN. Algoritme NWKNN dikembangkan seiring munculnya permasalahan data tidak seimbang pada data latih yaitu jumlah data latih antara satu kelas dengan kelas yang lain mempunyai selisih yang cukup besar. Perbedaan antara algoritme NWKNN dengan metode KNN adalah adanya pemberian bobot pada kelas/jenis yang berasal dari kategori mayoritas maka diberi nilai bobot kecil, sedangkan pada kategori minoritas akan diberi nilai bobot besar. Sehingga diharapkan dengan adanya bobot ini mengurangi kesalahan klasifikasi pada data uji yang menjadi anggota kelas minoritas.

Tahapan proses yang harus dilakukan untuk menerapkan algoritme NWKNN tidak jauh berbeda dengan langkah algoritme KNN, yang membedakan adalah perhitungan bobot dan score untuk menentukan klasifikasi terhadap data uji.

Perhitungan bobot dapat dilakukan sesuai dengan Persamaan 7:

$$Weight_i = \frac{1}{\left(\frac{Num(C_i^d)}{\text{Min}\{Num(C_n^d) | n=1, \dots, K^*\}} \right)^{1/exp}} \tag{7}$$

Keterangan

- $Num(C_i^d)$ = banyaknya data latih d pada kelas i
- $Num(C_j^d)$ = banyaknya data latih d pada kelas j , dimana j terdapat dalam himpunan k tetangga terdekat
- Exp = Eksponen (nilai exp lebih dari 1)

Setiap nilai bobot yang didapatkan akan digunakan untuk menghitung nilai skor data uji terhadap setiap kelas/jenis. Perhitungan skor pada metode NWKNN dapat dilakukan dengan Persamaan 8 dan 9.

$$Score(X, C_i) = Weight_i * \left(\sum_{d \in KNN(X)} \left(\left(\sqrt{\sum_{l=1}^n (x_{2l} - x_{1l})^2} \right) * \delta(d_j, C_i) \right) \right) \tag{8}$$

Atau

$$Skor(X, C_i) = Weight_i * \left(\sum_{d \in NWKNN(X)} ((Sim(q, d_j) * \delta(d_j, C_i))) \right) \tag{9}$$

Keterangan:

- $Weight_i$: Bobot jenis/kelas i
- $d \in NWKNN(X)$: data latih d_j pada kumpulan tetangga terdekat dari data uji X

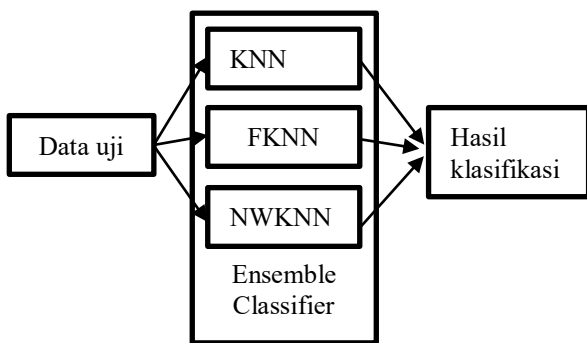
$$\sqrt{\sum_{l=1}^n (x_{2l} - x_{1l})^2}$$

: jarak antara data uji dan data latih

$\delta(d_j, C_i)$: akan bernilai 1 jika nilai jarak $\in C_i$ dan bernilai 0 jika nilai jarak $\notin C_i$
 $\text{Sim}(q, d_j)$: nilai *CosSim* antara data uji dan data latih.
 C_i : jenis atau kelas i

4. DESKRIPSI SISTEM

Sistem yang akan dibuat pada penelitian ini bertujuan untuk mendeteksi jenis Attention Deficit Hyperactivity Disorder (ADHD) pada anak usia dini dengan mengimplementasikan metode Ensemble Classifier dengan menggunakan 3 buah single classifier yaitu KNN, NWKNN dan F-KNN. Hasil dari pengklasifikasian ini yaitu tiga jenis ADHD yakni Hyperactivity, Impulsif, Inattention atau tidak ADHD. Input yang digunakan dalam sistem berupa jawaban yang dipilih dari setiap pernyataan dimana terdapat empat puluh lima pernyataan gejala yang nantinya akan merujuk kesalah satu jenis ADHD atau tidak termasuk ADHD dan kemudian jawaban dari inputan tersebut akan diubah kedalam nilai angka yang akan di proses dalam untuk diklasifikasikan. Nilai angka tersebut akan di proses oleh masing-masing *single classifier* sehingga setiap data yang di masukkan gejalanya akan menghasilkan output jenis ADHD. Apabila setiap data telah diproses oleh masing-masing *single classifier* maka hasil akhir jenis ADHD akan menjadi input di *ensemble classifier* sehingga akan di proses dengan *majority voting* atau penentuan berdasarkan klas mayoritas sehingga menghasilkan satu output jenis ADHD. Sebagai contoh apabila ada data gejala-gejala yang diinputkan oleh pengguna maka gejala tersebut akan diproses oleh *single classifier* terlebih dahulu. Apabila KNN mengeluarkan hasil klasifikasinya sebagai impulsif, FKNN sebagai inattention dan NWKNN sebagai impulsif maka apabila di proses oleh *ensemble classifier* maka hasil akhir klasifikasinya adalah impulsif karena klas mayoritas hasil *single classifier* adalah impulsif. Tahapan proses *Ensemble classifier* dapat dilihat pada Gambar 1



Gambar 1. Proses Ensemble Classifier

5. HASIL DAN PEMBAHASAN

Proses pengujian terhadap sistem dengan menggunakan metode KNN, FKNN, NWKNN dan

Ensemble Classifier. Proses pengujian tersebut meliputi pengaruh perubahan nilai K dan perubahan data latih terhadap data uji yang tetap dengan Cosine Similarity (Cossim). Pengujian terhadap perubahan nilai k digunakan untuk mendapatkan nilai k optimal yang dapat menghasilkan akurasi tertinggi, Sedangkan pengujian terhadap perubahan data latih dan data uji diperlukan untuk mengetahui rasio data latih dan data uji terbaik yang digunakan aplikasi untuk menghasilkan nilai akurasi terbaik. Pengujian dilakukan untuk mengetahui hasil akurasi dari implementasi metode yang telah dilakukan

5.1 Pengujian terhadap perubahan nilai K dengan menggunakan 60 data latih

Pengujian ini dilakukan dengan melakukan menggunakan 20 data uji 60 data latih. Pada pengujian ini di lakukan dengan menggunakan cossim untuk menghitung kedekatan diantara dua data. Hasil pengujian selengkapnya dapat dilihat pada Tabel 1

Tabel 1. Hasil Pengujian terhadap perubahan nilai K dengan menggunakan Cossim untuk data latih 60

No.	K	Akurasi (%)			
		KNN	NWKNN	FKNN	Ensemble Classifier
1.	2	75	70	80	80
2.	4	75	75	90	80
3.	8	75	65	80	75
4.	10	75	70	75	75
5.	15	55	55	55	55
6.	20	75	65	45	75
7.	25	60	65	45	60
8.	30	45	50	45	45
9.	35	45	45	45	45
10.	40	45	45	45	45
Rata-rata akurasi		62,5	60,5	60,5	63,5

Pada Tabel 1 dapat dilihat nilai rata-rata akurasi dari ensemble classifier lebih besar dibandingkan yang lain. Hal ini menunjukkan untuk data latih 60 akurasi ensemble classifier lebih baik di karenakan hasil klasifikasinya berdasarkan mayoritas klas yang dihasilkan single classifier.

5.2 Pengujian terhadap perubahan nilai K dengan menggunakan 70 data latih

Pengujian ini dilakukan dengan melakukan menggunakan 20 data uji 70 data latih. Pada pengujian ini di lakukan dengan menggunakan cossim untuk menghitung kedekatan diantara dua data. Hasil pengujian selengkapnya dapat dilihat pada Tabel 2.

Tabel 2. Hasil Pengujian terhadap perubahan nilai K dengan menggunakan Cossim untuk data latih 70

No.	K	Akurasi (%)			
		KNN	NWKNN	FKNN	Ensemble Classifier
1.	2	75	80	80	80
2.	4	85	85	90	85
3.	8	85	80	80	85

No.	K	Akurasi (%)			
		KNN	NWKNN	FKNN	Ensemble Classifier
4.	10	85	90	75	90
5.	15	90	90	55	90
6.	20	85	80	45	75
7.	25	65	70	45	65
8.	30	55	65	45	55
9.	35	45	50	45	45
10.	40	45	45	45	45
Rata-rata akurasi		71,5	73,5	60,5	71,5

Pada Tabel 2 dapat dilihat nilai rata-rata akurasi dari *ensemble classifier* lebih kecil dibandingkan NWKNN. Hal ini menunjukkan untuk data latih 70 akurasi *ensemble classifier* tidak lebih baik dari NWKNN di karenakan hasil klasifikasinya berdasarkan mayoritas klas yang dihasilkan *single classifier*. Kalau mayoritas klasnya salah klasifikasi maka hasil akhirnya juga salah klasifikasinya

5.3 Pengujian terhadap perubahan nilai K dengan menggunakan 80 data latih

Pengujian ini dilakukan dengan melakukan menggunakan 20 data uji 80 data latih. Pada pengujian ini di lakukan dengan menggunakan cossim untuk menghitung kedekatan diantara dua data. Hasil pengujian selengkapnya dapat dilihat pada Tabel 3

Tabel 3. Hasil Pengujian terhadap perubahan nilai K dengan menggunakan Cossim untuk data latih 80

No.	K	Akurasi (%)			
		KNN	NWKNN	FKNN	Ensemble Classifier
1.	2	80	80	80	80
2.	4	85	85	90	85
3.	8	95	90	80	95
4.	10	95	95	75	95
5.	15	95	80	55	95
6.	20	90	85	45	90
7.	25	70	75	45	70
8.	30	70	75	45	70
9.	35	55	60	45	55
10.	40	45	50	45	45
Rata-rata akurasi		78	77,5	60,5	78

Pada Tabel 3 dapat dilihat nilai rata-rata akurasi dari *ensemble classifier* lebih besar dibandingkan NWKNN dan FKNN tetapi sama besar dengan KNN. Hal ini menunjukkan untuk data latih 80 akurasi *ensemble classifier* sama dengan KNN karena mayoritas klas yang dihasilkan oleh *single classifier* sama dengan hasil klasifikasi menggunakan KNN.

5.4 Analisis hasil pengujian

Pengujian dengan mengubah nilai K menunjukkan bahwa nilai K optimal untuk beberapa percobaan yaitu pada K=10 dengan nilai akurasi maksimal yaitu 95%. Akan tetapi semakin besar nilai K yaitu diatas K=20 maka nilai akurasi untuk masing-masing algoritme akan semakin turun. Hal ini dikarenakan semua algoritme penentuan klasifikasinya

berdasarkan jumlah ketetanggannya. Maka semakin banyak jumlah tetangga yang di perhitungkan maka kemungkinan salah klasifikasinya semakin besar.

Pengujian dengan mengubah jumlah data latih menunjukkan bahwa semakin besar data latih maka akurasi semakin meningkat. Hal ini dikarenakan algoritme klasifikasi yang digunakan bergantung pada jumlah data latih sehingga semakin besar data latih kemungkinan untuk terjadinya kesalahan klasifikasi akan semakin sedikit.

Hasil perbandingan Ensemble Classifier dengan ketiga metode lain menghasilkan rata-rata akurasi yang baik. Meskipun hasil selisih rata-rata akurasi dengan metode lain tidak terlalu jauh hal ini menunjukkan bahwa *ensemble classifier* hasil akhirnya masih tergantung dari *single classifier* yang digunakan dikarenakan penentuan kelasnya berdasarkan *majority voting* yaitu penentuan klas berdasarkan hasil klas dari *single classifier* yang diambil klas mayoritasnya.

6. KESIMPULAN

Kesimpulan yang dapat diambil dari penelitian ini yaitu cara mengidentifikasi ADHD dengan menggunakan metode *Ensemble Classifier* yaitu dengan menggunakan hasil penentuan klas dari *single classifier* kemudian dengan menggunakan *majority voting* menentukan klas berdasarkan klas mayoritas. Hasil akurasi tertinggi dari *ensemble classifier* yaitu 95% dengan nilai K optimal yaitu K=10. Akan tetapi semakin besar nilai K yaitu diatas K=20 maka nilai akurasi untuk masing-masing algoritme akan semakin turun. Hal ini dikarenakan semua algoritme penentuan klasifikasinya berdasarkan jumlah ketetanggannya. Maka semakin banyak jumlah tetangga yang di perhitungkan maka kemungkinan salah klasifikasinya semakin besar.

DAFTAR PUSTAKA

- ADENIYI, D.A., WEI, Z., & YONGQUAN, Y.,2016. Automated web usage data mining and recommendationsystem using K-Nearest Neighbor (KNN)classification method. *Applied Computing and Informatic*
- GANG, WANG., dkk. 2013. Sentiment Classification: The Contribution of *Ensemble*.School of Management Hefei University of Technology. China.
- HURLOCK,E. 1980. *Psikologi Perkembangan Suatu pendekatan Sepanjang Rentang Kehidupan*,di Terjemahkan dari bahasa inggris oleh Penerbit Erlangga. Indonesia-Erlangga.
- JOWIK, A. 2013. *A Learning Scheme for A Fuzzy K-NN Rule*. Pattern Recognition Letters. Vol. 1, hal. 287-289.
- KUSNAWI. 2007. *Pengantar Solusi Data Mining. Seminar Nasional Teknologi 2007 (SNT 2007)*, Issue ISSN : 1978 – 9777

- NADIA FF DA SILVA, EDUARDO, R HRUSCHKA, ESTEVAM R HRUSCHKA. Tweet Sentiment analysis with classifier *Ensemble*. University of Sao Paulo, Sao Carlos Brazil.
- NASUTION, A.S. 2015. Informasi dan Teknologi Ilmiah (INTI): *Penerapan Algoritma Modified K-Nearest Neighbour(MKNN) Untuk Pengklasifikasian Penyakit Attentiondeficit Hiperactive Disorder (ADHD) PADA ANAK*. 2(5). 7-10.
- NEALE, J.M., KRING, A.M. 2004. *Psikologi Abnormal*. di Terjemahkan dari bahasa inggris oleh Fajar, Jakarta-Indonesia.
- 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).
- PUTRI, Y.C., ATASTINA, I., YULITA, I.N. 2012. *Analisis Dan Implementasi Teknik Fuzzy K-Nearest Neighbor Pada Kasus Imbalance Class*. Universitas Telkom.
- PRAMUDIONO, I., 2003. Pengantar Data mining: *Menambang Permata Pengetahuan di Gunung Data*. IlmuKomputer.com.
- TAN, SONGBO. 2005, "Neighbor-weighted K-nearest neighbor for unbalanced text corpus", *Expert Systems with Applications 28 (2005) 667-671*.
- VERBAETEN, SOFIE DAN ASSCHE, ANNELEEN VAN. 2003. *Ensemble Methods for Noise Elimination in Classification Problems*. T. Windeatt and F. Roli (Eds.): MCS 2003, LNCS 2709, pp. 317-325
- XIA, RUI., ZONG, CHENGQING., LI, SHOUSA. 2010. *Ensemble of Feature Set And Algorithm for Sentiment Classification*. National Laboratory and Pattern Recognition. China.

Halaman ini sengaja dikosongkan