
ALGORITMA K-NN DALAM MENGLASIFIKASIKAN DAERAH POTENSI PEMAIN BULUTANGKIS TUNGGAL PUTRA DI INDONESIA

Adi Nur Rohkhim ¹⁾, Muhammad Hafidh Al-Ghifari ²⁾, La Ode Muhammad Zulkifli Rinjani ³⁾, Natalia Damastuti ⁴⁾

^{1,2,3,4} Sistem Komputer, Fakultas Ilmu Komputer, Universitas Narotama Surabaya
Jln. Arief Rahman Hakim No.51 Surabaya

¹adinurrohkhim@gmail.com, ²Mh.alghifari007@gmail.com, ³oman.laode@gmail.com,

⁴natalia.damastuti@narotama.ac.id

ABSTRAK

Bulutangkis merupakan salah satu cabang olahraga yang dipertandingkan di ajang olimpiade musim panas. Tercatat pembulutangkis Indonesia telah memperoleh 7 medali emas di ajang tersebut. Hingga saat ini Indonesia belum bisa menambah medali emas satupun dari cabang olahraga lain yang dipertandingkan di Olimpiade. Tak heran bulutangkis menjadi olahraga yang sangat penting di Indonesia. Meskipun bulutangkis bukan berasal dari Indonesia, namun Indonesia telah melahirkan banyak legenda bulutangkis sejak tahun 1960-an hingga sekarang. Di era digital sekarang *sport science* telah di kembangkan di berbagai negara untuk mendukung kinerja atlet dan *official*, namun hal ini jika tidak di dukung dengan regenerasi pemain muda, maka estafet prestasi akan menjadi terputus. Betapa pentingnya mempersiapkan pemain muda berpotensi untuk meneruskan tradisi prestasi di cabang bulutangkis. Indonesia merupakan salah satu negara dengan regenerasi pemain muda yang cukup lambat dibanding negara kompetitor yang lain seperti China, Korea Selatan dan Jepang. Implementasi algoritma *K-Nearest Neighbor* untuk mengklasifikasikan wilayah dengan potensi atlet bulutangkis tunggal putra di Indonesia menjadi salah satu solusi agar induk organisasi bulutangkis Indonesia lebih muda mendapatkan pemain tunggal putra berpotensi. Dengan menggunakan 1000 data peringkat nasional tunggal putra di Indonesia dan mengklasifikasikannya menjadi 3 daerah yaitu daerah Berpotensi, cukup berpotensi, dan tidak berpotensi.

Kata kunci : K-Nearest Neighbor, klasifikasi, bulutangkis, Indonesia, data mining

ABSTRACT

Badminton is one sport that is contested at the Summer Olympics. Noted Indonesian badminton player has won 7 gold medals in the event. Until now Indonesia has not been able to add any gold medal from other sports that were competed in the Olympics. No wonder badminton has become a very important sport in Indonesia. Even though badminton is not from Indonesia, but Indonesia has given birth to many badminton legends since the 1960s until now. In the digital age sport science has now been developed in various countries to support athlete and official performance, but if this is not supported by the regeneration of young players, the achievement relay will be interrupted. How important it is to prepare young players with the potential to carry on the tradition of achievement in the badminton branch. Indonesia is one of the countries with a relatively slow regeneration of young players compared to other competing countries such as China, South Korea and Japan. The implementation of the K-Nearest Neighbor algorithm to classify areas with the potential of male singles badminton athletes in Indonesia is one solution so that the parent of Indonesian badminton organizations get a potential single male player. By using 1000 national male singles ranking data in Indonesia and classifying them into 3 regions, Potential areas, enough potential, and no potential.

Keywords: K-Nearest Neighbor, classification, badminton, Indonesia, data mining.

PENDAHULUAN

Indonesia adalah salah satu negara kepulauan terbesar di dunia. Dengan jumlah penduduk lebih dari 250 juta jiwa. Indonesia seharusnya lebih mudah mencari bakat bakat muda berpotensi di bidang olahraga khususnya cabang olahraga bulutangkis sektor tunggal putra. Karena Indonesia mempunyai materi dengan jumlah penduduk yang banyak yang tersebar di seluruh kepulauan Indonesia. Kenapa harus tunggal putra ? . Sejak kesuksesan Taufik Hidayat di ajang olimpiade tahun 2004 silam dengan menyumbangkan medali emas, sampai sekarang Indonesia belum berbicara banyak di sektor ini, bahkan medali perunggu pun pembulutangkis tunggal putra Indonesia belum bisa tercapai.

Sebenarnya ada beberapa faktor yang menyebabkan prestasi atlet bulutangkis sektor tunggal putra Indonesia mengalami penurunan. Dan salah satu faktor yang berperan penting adalah masalah mempersiapkan pemain muda sebagai regenerasi. Karena Indonesia mempunyai wilayah kepulauan yang cukup luas. Sehingga wajar jika pengurus besar persatuan bulutangkis Indonesia mempunyai kendala dalam mencari bibit muda di daerah. Mulai dari masalah waktu, SDM, hingga anggaran. Meskipun memang telah di bangun klub bulutangkis besar dengan fasilitas yang sangat mempuni, namun klub ini hanya terbatas di pulau Jawa.

Berdasarkan latar belakang tersebut maka diperlukannya suatu pengklasifikasian berdasarkan daerah dengan potensi pemain bulutangkis tunggal putra di Indonesia, dengan menggunakan salah satu metode pengklasifikasian di *data mining*. Yaitu metode KNN atau *K-Nearest Neighbor*. Daerah di Indonesia akan di bagi menjadi 3 wilayah, yaitu wilayah dengan potensi pemiaan bulutangkis sektor tunggal putra, wilayah dengan cukup potensi pemiaan bulutangkis tunggal putra dan wilayah tidak berpotensi. Tujuan dari penelitian ini adalah membantu kinejra dari tim pencari bakat dari klub maupun dari pengurus besar persatuan bulutangkis seluruh Indonesia dalam mendapatkan pemain muda berpotensi disektor tunggal putra.

Bulutangkis Indonesia

Di Indonesia sendiri, sejarah bulutangkis dimulai pada tahun 1930-an. Di masa itu cabang olahraga ini berada di bawah naungan organisasi ikatan *sport Indoneisa* (ISI). Namun bulutangkis sempat dilupakan karena Indonesia menghadapi masa perang. Namun pada tahun 1974 presiden Soekarno kampanye “nation building” sehingga bulutangkis kembali berkembang. (Vannisa 2018)

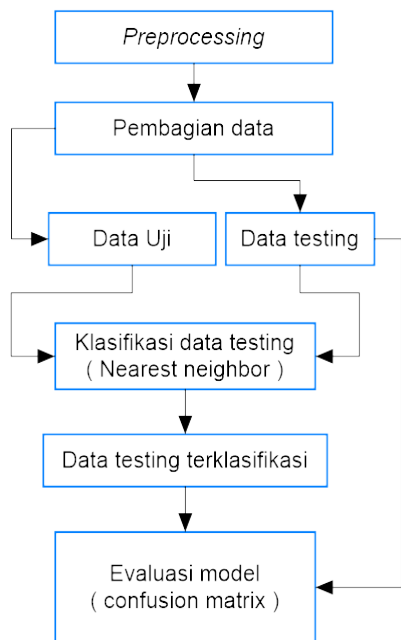
Sedangkan bulutangkis baru masuk menjadi salah satu cabang olahraga yang di pertandingkan di olimpiade pada tahun 1992 di Barcelona. Kala itu bulutangkis hanya mempertandingkan 4 sektor ,yaitu sektor tunggal putra , tunggal putri , ganda putri dan ganda putra. Tim bulutangkis Indonesia sejak itu tak pernah nihil mendapatkan medali emas kecuali pada tahun 2012 di London. Tim bulutangkis Indonesia gagal menyambung tali estafet tradisi emas. Namun hal ini terbayar lunas ketika tahun 2016 di Brazil. Indonesia berhasil menyumbangkan medali emas dari sektor ganda campuran atas nama Tantowi Ahmad dan Liliana Natsir. Di final mereka mengalahkan ganda campuran asal negeri jiran Malaysia dengan 2 game langsung .(DjarumFoundation 2012)

Klasifikasi

Klasifikasi adalah suatu kinerja yang mempunyai sebuah penilaian objek data untuk memasukkannya ke dalam suatu kelas tertentu dari beberapa kelas yang telah tersedia. Dalam proses pengklasifikasian tersebut terdapat dua proses yang harus di jalankan, yaitu yang pertama dengan membangun model data untuk disimpan sebagai memori dan patokan nilai , kemudian yang kedua menggunakan model tersebut untuk melakukan pengidentifikasian atau pengklasifikasian atau prediksi pada suatu data lain, dan hasil akhirnya adalah agar data tersebut bisa diketahui di kelas mana objek data tersebut dimasukkan berdasarkan model data yang

telah disimpan dalam memori pada proses sebelumnya. (Riyan Eko Putri, Suparti 2014)

Metode klasifikasi merupakan salah satu metode dari data Mining. Data mining itu sendiri adalah suatu istilah yang digunakan untuk menguraikan penemuan pengetahuan didalam database. Kinerja dari data mining yaitu memproses sebuah data dengan menggunakan teknik statistik, matematika, kecerdasan buatan, dan machine learning . Sehingga data mining berfungsi untuk mengekstrasi dan mengidentifikasi sebuah informasi yang bermanfaat dan pengetahuan yang terkait dari berbagai database yang mempunyai ukuran besar. (Sumarlin 2015)



Gambar 1. Pemodelan klasifikasi (Badu 2016)

K-Nearest Neighbor

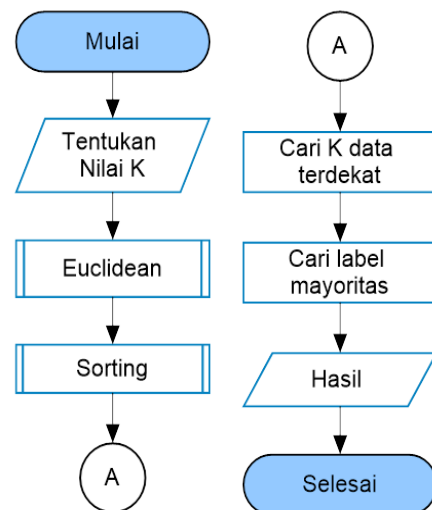
KNN adalah salah satu metode pengklasifikasian berdasarkan data *training* atau data pembelajaran dengan melihat dari posisi jarak yang paling dekat dengan objek berdasarkan nilai *k*. Metode ini bertujuan untuk mengklasifikasikan objek baru berdasarkan atribut dan training sample.

Objek diberikan suatu titik query, kemudiantitik tersebut akan ditemukan sejumlah *K* objek atau titik training yang paling dekat dengan titik query. Nilai prediksi dari query akan tersebut akan dihitung berdasarkan klasifikasi tetangga .

Sebelum melakukan perhitungan dengan metode *K-Nearest Neighbor*, terlebih dahulu harus menentukan data latih dan data uji. Kemudian akan dilakukan proses perhitungan untuk mencari jarak menggunakan *Euclidean*. Teknik ini sangat sederhana dan mudah diimplementasikan. Dengan mengelompokkan suatu data baru berdasarkan jarak data baru itu ke beberapa data/tetangga terdekat. Pertama sebelum mencari jarak data ke tetangga adalah menentukan nilai *K* tetangga (neighbor). Lalu, untuk mendefinisikan jarak antara dua titik yaitu titik pada data training dan titik pada data testing. (Nugraha and Faraby 2018)

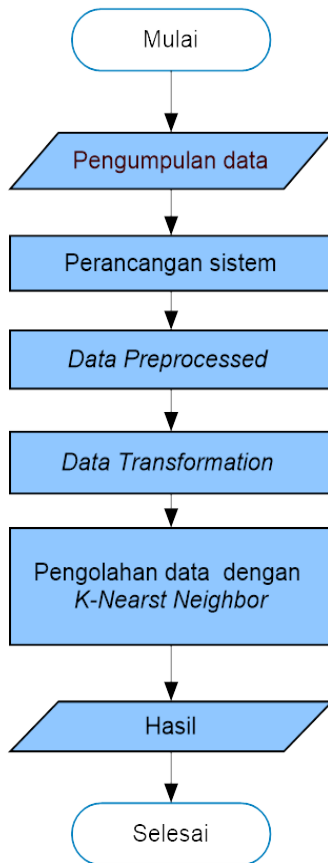
Algoritma *KNN* memiliki beberapa keunggulan yaitu kemampuan dalam mengatasi *training data* dengan banyak *noise* dan sangat efektif dalam menanagani *training data*-nya dengan data yang cukup besar. Sedangkan, kelemahan *KNN* yaitu *KNN* perlu menentukan nilai dari parameter *K* (jumlah dari tetangga terdekat). (Yustanti 2012)

Berikut flowchart metode *KNN*. (Nouvel 2015)



Gambar 2 .Flowchart *KNN*

METODE PENELITIAN



Gambar 3 .Flowchart metode penelitian

- **Pengumpulan data:** Pada tahap pengumpulan data, penulis menggunakan data primer dan data sekunder. Dimana data primer adalah data yang dihasilkan dari pengolahan sistem. Sedangkan data sekunder adalah data yang diperoleh dari website resmi organisasi persatuan bulutangkis seluruh Indonesia yaitu data daftar peringkat nasional bulutangkis sektor tunggal putra di Idnoneisa sebanyak 1000 data
- **Perancangan sistem:** Pada tahap perancangan sistem. Penulis melakukan perancangan pengklasifikasian data dengan 3 keluaran yaitu daerah dengan potensi pemain bulutangkis tunggal putra, daerah dengan cukup potensi

bulutangkis, dan daerah tidak berpotensi pemain bulutangkis tunggal putra

- **Data Preprocessed :** Pada tahap ini penulis menggunakan pembersihan data di beberapa atribut, karena pada data daftar peringkat nasional pemain bulutangkis tunggal putra terdapat atribut yang masih belum lengkap (*incomplete*) seperti pada atribut "Id BWF". Atribut ID BWF adalah ID yang diberikan oleh organisasi badan bulutangkis dunia pada pemain yang telah berkarir di dunia pembulutankisan dunia, atau bisa dikatakan pemain tersebut minimal pernah bermain atau berpartisipasi beberapa kali di turnamen kelas internasional sehingga mendapatkan point dan akhirnya BWF memberi ID pada pemain tersebut. Karena atribut ini tidak begitu berpengaruh pada penelitian ini. Akhirnya penulis melakukan pemersihan data. Agar data yang diolah bisa maksimal dan tidak mengalami eror
- **Transformation data:** Pada tahap ini. Penulis melakukan pengubahan tipe data pada beberapa atribut, agar pada saat data diolah dapat memenuhi asumsi asumsi yang mendasari analisa sistem. Serta penulis mengubah format file dari data yang diperoleh, yang awalnya file daftar peringkat nasional pemain tunggal putra bertipe ".exe", agar bisa diolah. Maka file tersebut di ubah formatnya menjadi "CSV"
- **Pengolahan data dengan KNN :** Pada tahap ini penulis menggunakan teknik pencarian tetangga terdekat yang umum dilakukan dengan menggunakan formula jarak *euclidean distance*. Jarak *euclidean distance* adalah formula untuk mencari jarak dengan 2 titik dalam ruang dua dimensi. Berikut rumus *euclidean*

$$d(x_i, x_j) = \sqrt{\sum_{r=1}^n (a_r(x_i) - a_r(x_j))^2} \quad (1)$$

Dimana :
 $d(x_i, x_j)$ = Jarak *Euclidean*
 $(x_i), (x_j)$ = *record ke-i, record ke-j*
 (a_r) = data ke r
 i, j = 1,2,3.....
 n = dimensi objek

- **Hasil** : Hasil adalah tahapan akhir dari penelitian ini. Setelah proses pengolahan klasifikasi dengan KNN selesai. Maka akan dihasilkan sebuah informasi untuk dasar pengembangan penelitian ini.

Pada data daftar peringkat nasional pemain bulutangkis tunggal putra, terdapat beberapa atribut bertipe nominal. Atribut tersebut akan diinisialisasikan menjadi bentuk numerik. Berikut inialisasi atribut "Wilayah"

Tabel 1. Inisialisasi wilayah

Inisialisasi	Wilayah	Jumlah
1	DKI Jakarta	313
2	Jawa Barat	148
3	Jawa Tengah	146
4	Jawa Timur	116
5	Banten	71
6	D.I. Yogyakarta	26
7	Kalimantan Selatan	13
8	Bali	7
9	Sulawesi Utara	13
10	Sumatera Selatan	22
11	Kalimantan Tengah	9
12	Sulawesi Selatan	11
13	Kalimantan Barat	5
14	Riau	8
15	Kalimantan Timur	32
16	Sumatera Barat	5
17	Nusa Tenggara Barat	6
18	Bengkulu	2
19	Kepulauan 14	7
20	Jambi	2
21	Sumatera Utara	24
22	Sulawesi Tenggara	2
23	Kepulauan Bangka Belitung	1
24	Papua	1
25	Gorontalo	5
26	Maluku Utara	1
27	Nanggroe Aceh Darussalam	2

28	Lampung	2
----	---------	---

Berikut tabel penentuan kelas klasifikasi berdasarkan ranking nasional pembulutangkis tunggal putra dan poin turnamen yang di dapat oleh setiap pemain

Tabel 2. Penentuan klasifikasi

Ranking (poin turnamen)	Frekuensi	Inisialisasi
1 – 333 (73.090 – 720)	333	Berpotensi
334 – 667 (720 – 162)	334	Cukup benpotensi
668 – 998 (160 -25)	333	Tidak berpotensi

HASIL DAN PEMBAHASAN

Tools yang digunakan pada penelitian ini adalah "Weka versi 3.8.4". Weka merupakan salah tools data mining yang mampu melakukan beberapa metode pengolahan data dengan cara yang sederhana, salah satunya adalah klasifikasi menggunakan metode *K-Nearest Neighbor*. Pengujian menggunakan data awal akan di ubah dengan susunan sebagai berikut :

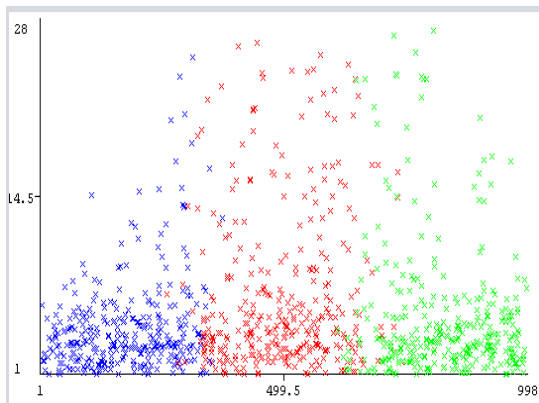
1. Jumlah data yang di gunakan adalah 1000 data yang di ambil dari web resmi organisasi persatuan bulutangkis republik Indonesia
2. Data awal mempunyai atribut sebanyak 7 atribut. Namun tidak semua atribut di gunakan untuk mengklasifikasikan. Hanya 4 atribut yang dipakai dalam penelitian ini.

Tabel 3. Data sampel daftar peringkat nasional

Rank-ing	poin	Wilayah	Klasifikasi
1	73090	1	Berpotensi
2	65680	2	Berpotensi
3	49000	2	Berpotensi

.	.	.	.
.	.	.	.
.	.	.	.
.	.	.	.
998	25	2	Tidak berpotensi

Tabel 3. adalah data sampel dari data daftar peringkat pemain bulutangkis tunggal putra di Indonesia.



Gambar 4. Visualisasi data klasifikasi daftar peringkat bulutangkis tunggal putra

Pada tahap pengujian peneliti menggunakan data uji sebanyak 400 data dan menggunakan 600 data sebagai data latih, dimana masing masing mempunyai 4 atribut yaitu atribut ranking nasional, poin turnamen, kode wilayah dan klasifikasi potensi.

Tabel 4. Data sampel daftar data latih

Rank-ing	poin	Wilayah	Klasifikasi
1	24600	1	Berpotensi
2	21440	3	Berpotensi
3	21290	2	Berpotensi
.	.	.	.
.	.	.	.
.	.	.	.
.	.	.	.
600	50	2	tidak berpotensi

Tabel 4. adalah daftar data sampel data latih yang berjumlah 600 data. Data latih adalah data yang sudah ada sebelumnya berdasarkan fakta dan yang sudah terjadi. Nantinya data latih ini akan bertugas melatih data uji.

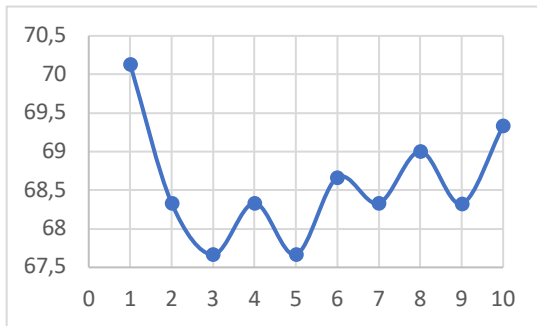
Tabel 5. Data sampel daftar data uji

Rank-ing	poin	Wilayah	Klasifikasi
1	16860	3	Berpotensi
2	16820	12	Cukup Berpotensi
.	.	.	.
.	.	.	.
.	.	.	.
.	.	.	.
400	45	2	tidak berpotensi

Tabel 5. adalah daftar data sampel data uji yang berjumlah 400 data. Data uji adalah data yang sudah berlabel yang tujuannya untuk menghitung nilai akurasi dari permodelan yang sudah dibentuk sebelumnya.

Tabel 6. Hasil pengujian

Kn	Akurasi (%)
K1	70.133
K2	68.333
K3	67.666
K4	68.333
K5	67.766
K6	68.666
K7	68.333
K8	69.003
K9	68.325
K10	69.333



Gambar 5. Grafik nilai akurasi

Dari hasil pengujian data testing terhadap data uji menggunakan KNN. Mulai dari K1 sampai dengan K10 didapatkan hasil. Nilai akurasi terendah terjadi pada K3 yaitu 67.666 % dan nilai akurasi tertinggi terjadi pada K1 yaitu 70.1333 %

PENUTUP

Setelah dilakukan klasifikasi daftar peringkat nasional pemain bulutangkis tunggal putra menggunakan metode KNN dapat disimpulkan bahwa bakat potensi pemain bulutangkis tunggal putra memang mayoritas terdapat di pulau Jawa. Namun hasil ini bukanlah patokan Organisasi persatuan bulutangkis Indonesia dalam mencari bibit pemain potensial tunggal putra di Indonesia.

DAFTAR PUSTAKA

- Badu, Zemi S. 2016. "PENERAPAN ALGORITMA K-NEAREST NEIGHBOR UNTUK KLASIFIKASI DANA DESA." (November).
- Djarum Foundation, Team. 2012. "Berharap Pada Mitos." 29 Juli 2012. <https://www.pbdjarum.org> (January 15, 2020).
- Nouvel, Ahmad. 2015. "Klasifikasi Kendaraan Roda Empat Berbasis Knn." 3(2): 66–69.
- Nugraha, Pratama Dwi, and Said Al Faraby. 2018. "Klasifikasi Dokumen Menggunakan Metode k -Nearest Neighbor (KNN) Dengan Information

Gain Document Classification Using k-Nearest Neighbor (k NN) Method with Information Gain." 5(1): 1541–50.

Riyan Eko Putri, Suparti, Rita Rahmawati. 2014. "Perbandingan Metode Klasifikasi Naïve Bayes Dan K-Nearest Neighbor Pada Analisis Data Status Kerja Di Kabupaten Demak Tahun 2012." *GAUSSIAN* 3: 831–38.

Sumarlin. 2015. "Implementasi Algoritma K-Nearest Neighbor Sebagai Pendukung Keputusan Klasifikasi Penerima Beasiswa PPA Dan BBM." 01: 52–62.

Vannisa. 2018. "Sejarah Bulu Tangkis." *June 24, 2018.* <https://perpustakaan.id> (January 15, 2020).

Tabel 4. Data sampel daftar data latih
2012. "Algoritma K-Nearest Neighbour Untuk Memprediksi Harga Jual Tanah." 9(1): 57–68.