

ALGORITMA K-NN GUNA KLASIFIKASI DAERAH POTENSI PEMAIN BULUTANGKIS TUNGGAL PUTRA DI INDONESIA

1Natalia Damastuti, 2Lukman Junaedi

*1,2Sistem Komputer, Fakultas Ilmu Komputer, Universitas Narotama, Jl. Arief Rahman Hakim No.51
Surabaya*

Email : natalia.damastuti@narotama.ac.id

*3Sistem Informasi, Fakultas Ilmu Komputer, Universitas Narotama, Jln Arief Rahman Hakim 51
Surabaya,*

Email : lukman.junaedi@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 Indonesai 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 atlit 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 pretasi 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 dapat digunakan untuk mengklasifikasikan wilayah dengan potensi atlit bulutangkis tunggal putra di Indonesia sehingga induk organisasi bulutangkis Indonesia lebih muda mendapatkan pemain tunggal putra berpotensi. Dengan menggunakan 1000 data peringkat nasional tunggal putra di Indonesia dan mengkalifikasikannya menjadi 3 daerah yaitu daerah Berpotensi , cukup berpotensi, dan tidak berpotensi dengan nilai K optimal pada K1 sebesar 70.133

Kata Kunci: *K-Nearest Neighbor, classification, badminton, Indonesia, data mining.*

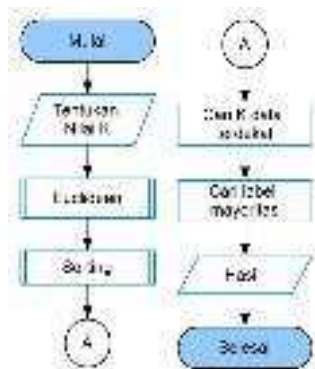
PENDAHULUAN

Indonesia adalah salah satu negara kepulauan terbesar di dunia dengan jumlah penduduk lebih dari 250 juta jiwa. Dengan jumlah penduduk sebesar tersebut seharusnya lebih mudah mencari bakat bakat muda berpotensi di bidang olahraga khususnya cabang olahraga bulutangkis sektor tunggal putra. Sejak kesuksesan pemain muda bulu tangkis, Taufik Hidayat, di ajang olimpiade tahun 2004 silam dengan menyumbangkan medali emas, sampai sekarang Indonesia belum berbicara banyak di sektor ini, bahkan medali perunggu tunggal putra Indonesia belum bisa tercapai. Sebenarnya ada beberapa faktor yang menyebabkan prestasi atlit 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. Selama ini penelitian yang menyangkut olah raga juga telah dilakukan beberapa peneliti dimana penelitian yang dilakukan mengenai motivasi dan variabel-variabel lain yang sesuai (Zeng et al. 2017). Berdasarkan latar belakang diatas, maka sangat diperlukan suatu klasifikasi daerah / wilayah di Indonesia berdasarkan potensi pemain bulutangkis. paper ini menyajikan suatu metode sederhana yang akan mengklasifikasikan daerah/wilayah potensi pemain muda yang ada di Indonesia, diharapkan dengan penelitian ini para pencari

bakat dapat lebih efektif dan efisien dalam menentukan wilayah paling banyak pemain muda berbakat.

Dalam penelitian ini dikembangkan teknik klasifikasi menggunakan metode klasifikasi *k*-NN. Pendekatan ini mengacu kepada beberapa penelitian tentang pengklasifikasian dengan beberapa pendekatan variabel yang berbeda-beda diantaranya adalah pengolahan citra digital yang dapat digunakan untuk klasifikasi mutu suatu buah dengan membedakan fitur-fitur citra (Prahudaya and Harjoko 2017). Pengenalan pemain dalam suatu olahraga juga dilakukan oleh (Foina et al. 2010) untuk mendeteksi pola pemain yang sedang bertanding. *k*-NN sendiri merupakan metode untuk melakukan klasifikasi terhadap obyek baru berdasarkan *k* tetangga terdekat dimana hasil dari query instance yang baru dapat diklasifikasikan berdasar maksimum dari kategori pada *k*-NN dan kelas yang paling banyak muncul akan menjadi kelas hasil klasifikasi (Zhai et al. 2015). Sementara itu metode ini juga dapat digunakan untuk klasifikasi dalam bidang medis. Klasifikasi jenis-jenis kanker telah dilakukan dengan metode ini. (Atthalla, Jovandy, and Habibie 2018) telah mengklasifikasikan penyakit dengan menggunakan perhitungan jarak kemiripan *minkowski*, yaitu jarak didalam ruang vektor yang telah ditentukan dan bisa dianggap sebagai generalisasi dari euclidean distance dan Manhattan (Atthalla, Jovandy, and Habibie 2018). Secara garis besar flowchart *k*-NN dapat disampaikan pada Gambar 1. (Nouvel 2015)



Gambar 1. Gambaran umum *k*-NN. Sumber : (Nouvel 2015)

METODE PENELITIAN

Dalam menyelesaikan suatu tujuan dari penelitian ini dilakukan tahapan-tahapan penelitian yang dapat diuraikan pada Gambar 2 dibawah ini.



Gambar 2. Tahapan Penelitian

Data dalam penelitian ini terdiri dari data primer dan data sekunder. Data primer merupakan data yang dihasilkan dari pengolahan sistem sedangkan data sekunder didapat dari website resmi organisasi persatuan bulutangkis dimana data yang terkumpul sebanyak 1000 data untuk sektor tunggal putra di seluruh Indonesia. 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. Tahap selanjutnya merupakan data preprocessing dimana akan dilakukan data cleaning di beberapa atribut seperti atribut ID BWF adalah ID yang diberikan oleh organisasi badan bulutangkis dunia pada pemain yang

telah berkarir di dunia pembulutankisan dunia, atau bisa di katakan 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 error. Transformasi data dilakukan untuk mengubah tipe dari atribut data dengan format 'exe' menjadi bentuk 'csv' agar dapat diolah lebih mudah.

Pada proses pengolahan data menggunakan k-NN teknik pencarian tetangga terdekat yang umum dilakukan dengan menggunakan formula jarak euclidean distance. Jarak euclidean dstance adalah formula untuk mencari jarak dengan 2 titik dalam ruang dua dimensi. Berikut rumus euclidean

$$d_{ij} = \sqrt{(x_i - x_j)^2 + (y_i - y_j)^2 + \dots + (z_i - z_j)^2} \tag{1}$$

Dimana :

- d_{ij} = Jarak Euclidean
- i, j = record ke-i, record ke-j
- x, y, z, \dots = data ke r
- i, j = 1,2,3,.....
- n = dimensi objek

Dalam tahapan 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 inisialai atribut "Wilayah". Terdapat beberapa inialisasi dapat disampaikan pada Tabel 1 dibawah ini. Berdasarkan tabel tersebut terdapat 28 inialisasi wilayah di seluruh Jawa Timur. Data tersebut merupakan peringkat pemain dalam tingkat Nasional. Seperti DKI Jakarta dengan jumlah 313, Jawa Barat

berjumlah 148, Jawa Timur memiliki nilai 116 dan beberapa wilayah lainnya.

Tabel 1. Tabel inialisasi Wilayah

Inisia-lisasi	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

Tahap klasifikasi dibagi menjadi 3 klaster sebagaimana data pada Tabel 2 dibawah. Penentuan kelas klasifikasi berdasarkan rangking nasional pembulutangkis tunggal putra dan poin turnamen yang di dapat oleh setiap pemain. Terdapat 3 klasifikasi wilayah yaitu wilayah dengan kategori 'Berpotensi

dengan point turnamen 1 – 333, wilayah ‘Cukup Berpotensi’ dengan point turnamen 334-667 dan wilayah dengan kategori ‘Tidak Berpotensi’ dengan nilai point turnamen 668 – 998 sebagaimana dapat di uraikan pada Tabel 2.

Tabel 2. Penentuan Klasifikasi

Rangking (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

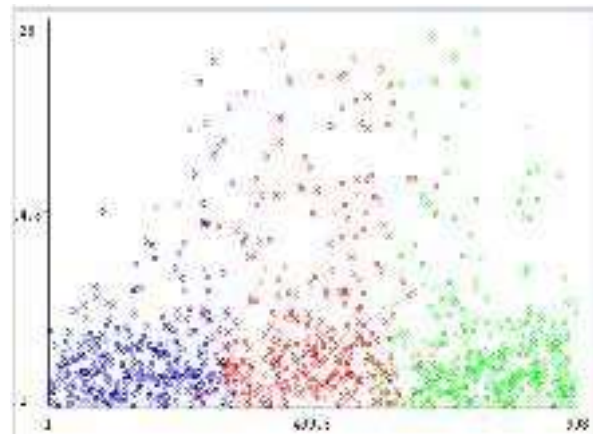
Pengolahan data yang digunakan pada penelitian ini adalah “Weka versi 3.8.4”. Weka merupakan salah tools data mining yang mampu melakukan beberapa metode pengolahan data dengan cara yang sederhana, salah satunya adalah klasifikasi menggunakan metode K-Nearest Neighbor. Pengujian menggunakan data awal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

Jumlah data yang di gunakan adalah 1000 data yang di ambil dari web resmi organisasi persatuan bulutangkis republik Indonesia, Data awal mempunyai atribut sebanyak 7 atribut. Namun tidak semua atribut di gunakan untuk mengklasifikasikan. Hanya 4 atribut yang dipakai dalam penelitian ini yaitu ‘Ranking’, ‘Poin’, ‘Wilayah’, dan ‘Klasifikasi’. Tabel 3 merupakan hasil pengolahan data yang didapat, dimana jika dilakukan perankingan maka ranking tertinggi memiliki poin sebesar 7.3090 dimana hasil ini berada di wilayah 1 sehingga dapat di klasifikasikan sebagai daerah yang memiliki potensi tinggi yang memiliki bibit-bibit pemain muda bulu tangkis di wilayah Indonesia. Rangking berikutnya memiki point permainan sebesar 65.680 yang berada di wilayah 2 dan masih diklasifikasikan sebagai wilayah yang memiliki potensi tinggi pemain-pemain tunggal putra. Sedangkan peringkat terakhir ada pada rangking 998

dimana hanya memiliki poin sebesar 25 dan berada pada wilayah 2 sehingga berdasarkan hasil poin tersebut maka wilayah ini dapat dikatakan tidak memiliki potensi pemain tunggal putra yang bagus dalam permainan hal ini dapat dilihat dari point yang dimiliki.

Tabel 3. Tabel 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



Gambar 4. Sebaran Data Klasifikasi

Gambar 4 menunjukkan sebaran data klasifikasi berdasarkan 4 atribut yang digunakan.

Pada tahap pengujian digunakan 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 menunjukkan

contoh data latih yang digunakan. Sebanyak 600 data atau 60% digunakan sebagai data latih dan sisanya menjadi data uji.

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 belatih data uji.

Tabel 5. Data sampel daftar data uji

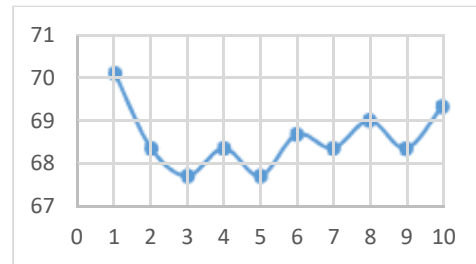
Rank-ing	poin	Wilayah	Klasifikasi
1	16860	3	Berpotensi
2	16820	12	Cukup Berpotensi
.	.	.	.
.	.	.	.
.	.	.	.
.	.	.	.
400	45	2	tidak berpotensi

Tabel 5. adalah daftar data sampel data uji yang berjumlah 400 data atau sebesar 40%. Data uji adalah data yang sudah berlabel dengan tujuannya untuk menghitung nilai akurasi dari permodelan yang sudah dibentuk sebelumnya.

Berdasarkan perhitungan *k*-nearest neighbor, didapatkan nilai akurasi sebagaimana ditampilkan dalam Tabel 6 dibawah.

Tabel 6. Hasil pengujian

K	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 latih 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

Berdasarkan hasil pengolahan dan analisa data, didapatkan bahwa daftar peringkat nasional pemain bulutangkis tunggal putra menggunakan metode KNN dapat di simpulkan bahwa bakat potensi pemain bulutangkis tunggal putra memang mayoritas terdapat di pulau Jawa. Algoritma *k*-NN memiliki tingkat akurasi yang dapat digunakan dengan *k* maksimal adalah *k*=1 dengan nilai akurasi sebesar 70.133

DAFTAR PUSTAKA

- Atthalla, Ikhsan Nuh, Adithia Jovandy, and Hanif Habibie. 2018. "Klasifikasi Penyakit Kanker Payudara Menggunakan Metode K-Nearest Neighbor." Prosiding Annual Research Seminar 4(1): 978-79.*
- Foina, Aislan Gomide, Rosa M. Badia, Ahmed El-Deeb, and Francisco Javier Ramirez-Fernandez. 2010. "Player Tracker - A Tool to Analyze Sport Players Using RFID." 2010 8th IEEE International Conference on Pervasive Computing and Communications Workshops, PERCOM Workshops 2010: 772-75.*
- Nouvel, Ahmad. 2015. "Klasifikasi Kendaraan Roda Empat Berbasis Knn." Bianglala Informatika 3(2): 66-69.*
- Prahudaya, Taftyani Yusuf, and Agus Harjoko. 2017. "Metode Klasifikasi Mutu Jambu Biji Menggunakan Knn Berdasarkan Fitur Warna Dan Tekstur." Jurnal Teknosains 6(2): 113.*
- Zeng, Howard Z., Wen-Yan Meng, Peng Sun, and Li-Sheng Xie. 2017. "A Study of Youth Badminton Players' Participation Motivation and Relate Elements*." IOSR Journal of Sports and Physical Education 04(03): 44-54.*
- Zhai, Zhenzhen, Huiqin Jiang, Liping Lu, and Yumin Liu. 2015. "Adaptive Truncation Coding for Computed Tomography Images." Proceedings of the 2014 International Symposium on Information Technology, ISIT 2014 02(1): 115-18.*