

Penerapan Algoritma K-Nn untuk Klasifikasi Gamers Usia Sekolah

Billyanto Satrian^a, Gusrianty^b

^a Sekolah Tinggi Ilmu Komputer Pelita Indonesia, Jl. Ahmad Yani No. 78-88, Pekanbaru, billyantosatria@student.pelitaindonesia.ac.id

^b Sekolah Tinggi Ilmu Komputer Pelita Indonesia Jl. Ahmad Yani No. 78-88, Pekanbaru, gusrianty@lecturer.pelitaindonesia.ac.id

INFORMASI ARTIKEL

Sejarah Artikel:

Diterima Redaksi: 11 Maret 2020

Revisi Akhir: 4 April 2020

Diterbitkan Online: 30 April 2020

KATA KUNCI

Gamers, Student, Score Grade Decrease,

K-Nearest Neighbor, School

KORESPONDENSI

E-mail:

gusrianty@lecturer.pelitaindonesia.ac.id

ABSTRACT

Di zaman yang serba canggih sekarang ini, kita pasti tidak asing lagi dengan yang namanya *smartphone* atau *game mobile*. Hampir di semua kalangan usia saat ini dapat memainkan *game mobile* yang terdiri dari berbagai macam jenis permainan. Di kalangan anak-anak usia sekolah di lingkungan Sekolah Kristen Kalam Kudus, kecanduan terhadap *game online* telah mulai mempengaruhi pencapaian akademik. Dampak negatif yang terjadi pada Siswa-siswi sekolah Kristen Kalam Kudus lebih banyak dari pada dampak positif. Untuk itu perlu dilakukan klasifikasi terhadap siswa sekolah Kristen Kalam Kudus untuk mengetahui sejauh mana *game online* turut memberi pengaruh terhadap keseharian dan pencapaian prestasi akademik siswa. Untuk mengetahui itu, ilmu data mining sangat cocok digunakan. Data mining merupakan sebuah ilmu dasar yang dapat membantu melahirkan informasi dari data yang ada. Algoritma klasifikasi k-Nearest Neighbor (k-NN) merupakan salah satu algoritma yang terdapat dalam data mining.

1. PENDAHULUAN

Di zaman yang serba canggih sekarang ini, kita pasti tidak asing lagi dengan yang namanya *game mobile* atau *smartphone*. Hampir di semua kalangan usia memainkan *game mobile* yang terdiri dari berbagai macam jenis permainan. Sebutan untuk setiap orang yang bermain dan menyukai *game* adalah *gamers*.

Kehadiran video *game* di tengah-tengah laju teknologi telah membawa pengaruh besar terhadap perkembangan pribadi dan adaptasi remaja, bahkan tidak sedikit remaja yang berubah menjadi pecandu *game* sehingga lupa pada jati diri mereka yang sesungguhnya.

Penelitian ini dilakukan di Sekolah Kalam Kudus Pekanbaru. Penelitian dilakukan untuk mengklasifikasikan dan mengetahui dampak negatif dalam bermain *game online* secara berlebihan, dan sejauh mana *game online* turut memberi pengaruh terhadap keseharian dan pencapaian prestasi akademik siswa. Permasalahan klasifikasi ini bisa diselesaikan dengan beragam metode, salah satunya adalah *K-Nearest Neighbor* (kNN). Algoritma klasifikasi *k-Nearest Neighbor* (k-NN) merupakan salah satu algoritma yang terdapat dalam data mining. k-NN adalah algoritma yang berfungsi untuk melakukan klasifikasi suatu data berdasarkan data pembelajaran (*train data*

sets), yang diambil dari tetangga terdekatnya (*nearest neighbors*).

k-NN melakukan klasifikasi dengan proyeksi data pembelajaran pada ruang berdimensi banyak. Ruang ini dibagi menjadi bagian-bagian yang merepresentasikan kriteria data pembelajaran. Setiap data pembelajaran direpresentasikan menjadi titik-titik pada ruang dimensi banyak. Berbagai penelitian tentang k-NN telah banyak dilakukan.

Pada penelitian sebelumnya telah membahas tentang klasifikasi untuk mengetahui dampak negatif dalam bermain *game online* secara berlebihan, algoritma k-NN digunakan untuk mengelompokkan data pemakaian *game online* secara berlebihan [1]. Pada Penelitian lain klasifikasi untuk mengetahui karakteristik pelajar yang bermain *game online* dikalangan pelajar yang dibuat oleh [2]. Pada penelitian ini digunakan data siswa-siswi sekolah Kristen Kalam Kudus pengguna *game online* yang diperoleh dari survey.

Data dipersiapkan dengan mengikuti langkah persiapan data dalam Knowledge Discovery in Database (KDD). Untuk kebutuhan training menggunakan kriteria yang terdiri dari usia, lama bermain, dan dampak negatif yang di timbulkan untuk mempresentasikan klasifikasi yang diinginkan [3].

2. TINJAUAN PUSTAKA

2.1 Gamers

Gamers adalah seseorang yang bermain dan menyukai game. Tidak peduli jika dia hanya bermain satu game saja, tidak memiliki peralatan yang berkualitas tinggi, bahkan jika dia tidak terlalu hebat dalam memainkan game tersebut. Selama dia masih bermain dan menyukai game maka orang tersebut adalah *gamer*.

2.2 Teknik Klasifikasi Dalam Data Mining

Teknik klasifikasi dalam data mining kelas banyak dilakukan. Pada penelitian[4], teknik klasifikasi yang digunakan adalah decision tree untuk menentukan kriteria mahasiswa dalam memilih program studi. Teknik decision tree juga ditemui dalam penelitian[5] yang memadukannya dengan algoritma fuzzy genetic untuk memprediksi hasil akademik siswa. Teknik klasifikasi lain yang diterapkan adalah Naive Bayes dan Random Forest[6]. Kedua teknik dimanfaatkan untuk menganalisis barang laris dan tidak. Selain barang, Naive Bayes dan Random Forest juga berhasil mendeteksi kecurangan dalam penggunaan kartu kredit [7], serta mengelompokkan gambar multispectral dalam data set yang besar[8].

2.2.1 Metode Algoritma K-NN (Nearest Neighbor)

K-Nearest Neighbor merupakan salah satu metode yang digunakan dalam pengklasifikasian. Prinsip kerja *K-Nearest Neighbor* (KNN) adalah mencari jarak kerja terdekat antara data yang akan dievaluasi dengan K tetangga (*Neighbor*) terdekatnya dalam data pelatihan. Berikut rumus pencarian jarak menggunakan *euclidian*[9].

Sedangkan Menurut Mariana (DINAMIKA INFORMATIKA – Vol.7 No. 1, Maret 2015, ISSN 2085-3343) Algoritma k-Nearest Neighbor atau k-NN adalah sebuah algoritma untuk melakukan klasifikasi terhadap objek berdasarkan data pembelajaran yang jaraknya paling dekat dengan objek tersebut. Tujuan dari algoritma KNN adalah mengklasifikasikan objek baru berdasarkan atribut dan sampel-sampel dari data training. Menurut Jollyta (2018) didalam jurnal *Decision Tree and Chi-Square Analysis to Determine Student's Criteria Choosing Study Program* (American Journal of Engineering Research (AJER) e-ISSN: 2320-0847 p-ISSN : 2320-0936 Volume-7, Issue-10) langkah - langkah persiapan untuk data didalam Knowledge Discovery in Database (KDD) adalah :

- a. *Selection*
Kegiatan yang dilakukan adalah pemilihan atau segmentasi data berdasarkan kriteria tertentu. Dalam kegiatan ini, data dipilih dari banyak data untuk dipilih kembali.
- b. *Pre-processing*
Pada tahap ini, pembersihan data dilakukan, di mana *field* yang tidak mendukung dapat dibuang. Selain itu data akan dikonfigurasi ulang untuk memastikan format tetap konsisten.
- c. *Transformation*
Kegiatan yang mentransformasikan data sehingga data dapat digunakan dan dilacak. Dengan kata lain, transformasi melakukan pemetaan data yang kompleks

Cara kerja perhitungan Algoritma *K-Nearest Neighbor* yaitu :

1. Menentukan parameter K sebagai banyaknya jumlah tetangga terdekat dengan objek baru.
2. Menghitung jarak antar objek/data baru terhadap semua objek/data yang telah di training.
3. Urutkan hasil perhitungan tersebut.
4. Tentukan tetangga terdekat berdasarkan jarak minimum ke K.
5. Tentukan kategori dari tetangga terdekat dengan objek/data.
6. Gunakan kategori mayoritas sebagai klasifikasi objek/data baru.
7. Selesai

2.3 Game Online

Game online adalah sebuah jenis video permainan yang hanya dapat dijalankan apabila suatu perangkat yang digunakan untuk bermain game terhubung dengan jaringan internet. Berkaitan dengan hal tersebut Burhan dan Tsharir (2005) mengemukakan bahwa game online merupakan permainan komputer yang dapat dimainkan oleh banyak pemain melalui internet. Kalau tidak, maka besar kemungkinan game online itu tak dapat dimainkan[10].

2.4 Sekolah

Sekolah adalah lembaga atau tempat yang dirancang khusus dimana anak-anak yang berusia 6-17 tahun belajar menuntut ilmu di bawah pengawasan guru. Banyak ahli menganggap masa ini sebagai masa tenang atau masa *latent*, di mana apa yang telah terjadi dan dipupuk pada masa-masa sebelumnya akan berlangsung terus untuk perkembangan masa-masa selanjutnya. Perkembangan dapat diartikan sebagai proses perubahan kuantitatif dan kualitatif individu dalam rentang kehidupannya, mulai dari masa konsepsi, bayi, kanak-kanak, masa remaja, sampai dengan dewasa [2].

3. METODOLOGI

Salah satu dari metode penelitian yang dimaksud adalah *System Development Life Cycle* (SDLC), dengan skema tahapan dimulai dari tahapan pengumpulan data, analisis data, perancangan sistem, implementasi sistem, pengujian sistem dan penyimpulan.

Adapun metode penelitian yang digunakan oleh penulis pada sistem dengan menggunakan Algoritma *k-Nearest Neighbor*. Alur tahapan *k-Nearest Neighbor* adalah penentuan nilai k, menghitung antar objek/data yang telah di training, urutkan hasil perhitungan, tentukan tetangga terdekat berdasarkan jarak minimum ke k, tentukan kategori tetangga terdekat dengan objek/data, gunakan kategori mayoritas sebagai klasifikasi objek/data baru tersebut.

4. HASIL DAN PEMBAHASAN

4.1. Data Penelitian

4.1.1. Data Training

Data penelitian diperoleh dari penyebaran 300 buah angket kepada siswa Sekolah Menengah Pertama (SMP) dengan jumlah siswa sebanyak 291 orang dan Sekolah Menengah Atas (SMA) sebanyak 257 orang di sekolah Kristen Kalam Kudus Pekanbaru dengan menggunakan google form. Jumlah sebaran dilakukan secara acak (random). Angket yang diisi oleh siswa SMP sejumlah 100 buah dan SMA 100 buah. Angket terdiri dari empat pertanyaan dimana pertanyaan tersebut menjadi kriteria yang menjadi dasar klasifikasi. Kriteria tersebut adalah umur, total pembelian voucher, lama bermain dan level tertinggi.

Tabel 1 Data Training

No	Nama	Umur	Beli Voucher	Lama bermain	Level	Klasifikasi
1	Gu	14	2	2	60	Ringan
2	Ry	14	4	2	70	Ringan
3	Di	14	0	1	65	Normal
4	Ro	14	8	3	44	Ringan
5	Gi	13	7	4	55	Sedang
6	Je	14	0	5	66	Berat
7	Fe	14	2	6	77	Berat
8	Oc	14	3	8	89	Berat
9	Se	14	0	9	90	Berat
10	Ga	14	9	4	99	Berat
...	...	...	...	...	...	...
200	Sa	18	10	7	60	Berat

4.1.2. Data Testing

Kemudian pada tahap selanjutnya diberikan data Testing yang bertujuan untuk mengetahui hasil klasifikasi yang akan didapatkan dengan cara menggunakan metode k-NN

Tabel 2 Data Testing

Nama	Umur	Total	Lama bermain	Level	Klasifikasi
Bu	16	10	5	60	??

Penggunaan data testing diatas bertujuan untuk mengetahui pola yang sesuai dengan data training.

4.2. Penggunaan Algoritma K-Nearest Neighbor

4.2.1. Tentukan Nilai K

Setelah itu nilai parameter yang dipilih akan menyimpulkan data testing yang akan dicari kelasnya tersebut berdasarkan nilai k data training terdekat dengan data testing. Pada kasus ini nilai parameter k nya adalah 55.

4.2.2. Perhitungan Euclidean Distance

Penentuan pola ditentukan melalui proses perhitungan data training menggunakan teknik euclidean distance untuk mendapatkan nilai terdekat. Berikut ini adalah rumus euclidean distance.

$$d = \sqrt{(q_1 - p_1)^2 + (q_2 - p_2)^2 + \dots + (q_n - p_n)^2}$$

Dimana :

d = jarak antara data training dan data testing

P = Kasus baru

q = Kasus yang ada dalam penyimpanan

n = Jumlah atribut dalam tiap kasus

Selanjutnya setelah mendapatkan hasil dengan rumus euclidean distance tabel dibawah berfungsi untuk menentukan hasil klasifikasi ketika sudah mendapatkan nilai dari perhitungan euclidean distance.

Tabel 3 Standar Nilai Klasifikasi

Klasifikasi	
Nilai Kategori	Jenis Kategori
1-50	Normal
51-100	Ringan
101-150	Sedang
151-200	Buruk

Salah satu nilai yang akan diambil yaitu dari data nomor 1 Gustian dan juga data testing pada tabel 4.2 yang sudah disediakan

$$d = \sqrt{(14 - 16)^2 + (2 - 10)^2 + (2 - 5)^2 + (60 - 60)^2}$$

$$d = 5.3851$$

Berikut ini adalah perhitungan manual euclidean distance

Tabel 4 Perhitungan Euclidean Distance

No	Ranking	Euclidean Distance	Klasifikasi
1	14	8.775	Ringan
2	42	12.2066	Ringan
3	40	12.0416	Normal
4	62	16.3707	Ringan
5	8	6.6332	Sedang
6	38	11.8322	Berat
7	74	18.9209	Berat
8	115	30.05	Berat
9	126	31.9374	Berat
10	150	39.0768	Berat
...	...	...	...
200	1	2.8284	Berat

4.2.3. Sortir Euclidean Distance

Setelah mendapatkan hasil perhitungan *euclidean distance*, pada tahap ini akan disortir nilai *euclidean distance* dari yang terkecil hingga terbesar dengan ketentuan $k=55$, untuk menentukan jenis kategori yang paling banyak muncul sehingga dapat menentukan klasifikasi pada data testing.

Tabel 5 Sortir *Euclidean Distance*

No	Ranking	Euclidean Distance	Klasifikasi
200	1	3.4641	Sedang
50	2	3.1623	Sedang
38	3	2.8284	Berat
51	4	4.1231	Berat
20	5	5.7446	Ringan
179	6	5.831	Ringan
131	7	6.1644	Berat
5	8	6.6332	Sedang
63	9	7.2111	Normal
22	10	7.8102	Ringan
...	...	...	...
91	55	13.9284	Berat

4.2.4. Menetapkan Kelas

Setelah data training diurutkan berdasarkan *euclidean distance* terkecil, selanjutnya dipilih 55 data terdekat (karena nilai $k=55$) antara data training dan data testing. Kemudian data *Euclidean Distance* yang telah di sortir dari tabel 4.5 dapat kita tentukan bahwa klasifikasi dari tetangga terdekat yang paling banyak muncul adalah klasifikasi di tingkat ringan.

Tabel 6 Hasil Klasifikasi

Berat	14
Sedang	14
Ringan	18
Normal	9

Tabel 7 Detail Hasil

Nama	Umur	Total	Lama bermain	Level	Klasifikasi
Bu	16	10	5	60	Ringan

Maka dapat dilihat pada Tabel 4.4 klasifikasi yang muncul pada data training lebih dominan nilai kategori yang ringan. Maka disimpulkan bahwa data testing yang diuji masuk kedalam klasifikasi ringan.

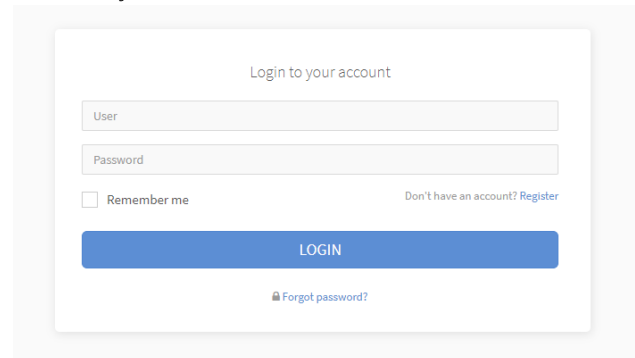
4.3. Antarmuka Sistem

Tampilan antarmuka dari sistem ini diharapkan dapat memberikan tampilan konsep yang sudah diimplementasikan dari ide dan rancangan pada tahap perancangan sistem sehingga

dapat menjadi acuan dan memudahkan bagi pengguna untuk memahami cara kerja, penggunaan sistem dan fitur-fitur yang tersedia pada sistem.

1. Halaman Login

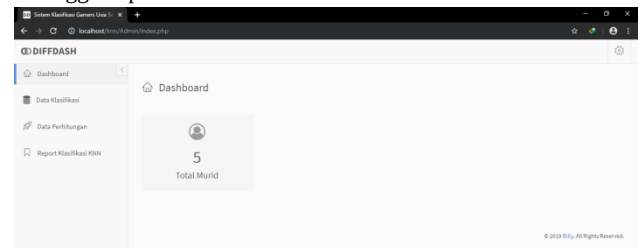
Halaman Login ini digunakan sebagai halaman dasar yang bertujuan untuk admin/user masuk ke halaman utama.



Gambar 1. Tampilan Halaman Login

2. Halaman Dashboard

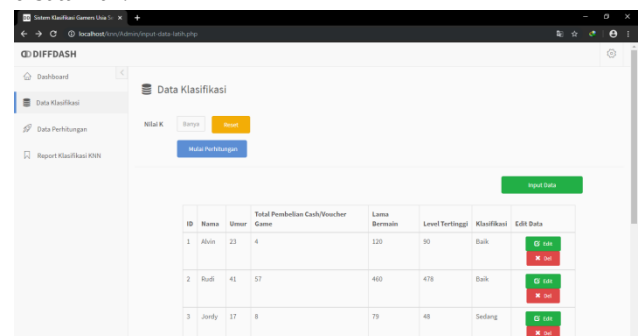
Halaman yang berisikan fitur-fitur yang sudah disediakan sehingga dapat diakses oleh admin.



Gambar 2. Tampilan Halaman Dashboard

3. Tampilan Halaman Data Klasifikasi

Tampilan Halaman Data Klasifikasi digunakan untuk menginput data dan juga memasukkan nilai k yang dibutuhkan.



Gambar 3. Tampilan Halaman Data Klasifikasi

4. Tampilan Halaman Input Data

Menampilkan halaman input data, jika ada data yang baru atau ingin diuji.

Gambar 4. Tampilan Halaman Input Data

5. Halaman Hasil Data Perhitungan
Menampilkan hasil data yang sudah dihitung dan kemudian dibandingkan.

Gambar 5. Tampilan Halaman Hasil Data Perhitungan

6. Halaman Hasil Klasifikasi
Halaman yang berisikan hasil akhir dari proses perhitungan data.

Gambar 6. Tampilan Halaman Hasil Klasifikasi.

5. KESIMPULAN DAN SARAN

Berdasarkan hasil penelitian yang telah dilakukan, maka dapat disimpulkan sebagai berikut :

Terbukti dengan menggunakan sistem berbasis web dengan metode K-nn dalam mengklasifikasi data siswa menjadi lebih cepat dan mudah, serta dapat membantu para guru di sekolah membimbing siswa yang kecanduan game supaya bisa berubah ke arah yang lebih baik.

Berdasarkan penelitian yang dilakukan, maka terdapat beberapa saran sebagai berikut :

1. Menyarankan kepada guru agar dapat menggunakan sistem berbasis web secara baik sehingga dapat membantu siswa yang kecanduan bermain game.
2. Bagi peneliti selanjutnya jika ingin mengembangkan penelitian ini dapat menambahkan atau menggunakan metode lain, seperti metode SPK sehingga dapat menentukan dampak yang akan ditimbulkan jika bermain game secara berlebihan .

UCAPAN TERIMA KASIH

Ucapan terima kasih setulus-tulusnya, penulis haturkan kepada Institut Bisnis dan Teknologi Pelita Indonesia dan seluruh pihak yang berkontribusi dan membantu dalam penelitian ini.

DAFTAR PUSTAKA

- [1] D. Edy Kurniawan Program Studi Bimbingan dan Konseling, “Pengaruh Intensitas Bermain Game Online Terhadap Perilaku Prokrastinasi Akademik Pada Mahasiswa Bimbingan Dan Konseling Universitas PGRI Yogyakarta,” *J. Konseling GUSJIGANG*, vol. 3, no. 1, p. 98, 2017.
- [2] R. N. Nasution, “DAMPAK POSITIF DAN NEGATIF PERMAINAN GAME ONLINE DIKALANGAN PELAJAR,” vol. 3, no. 2, pp. 54–67, 2015.
- [3] M. Ulfa, “Pengaruh Kecanduan Game Online Terhadap Perilaku Remaja di Mabes Game Center Jalan HR.Subrantas Kecamatan Tampan Pekanbaru,” *Jom . Fisip*, vol. 4, no. Pengaruh Kecanduan Game Online, pp. 1–13, 2017.
- [4] D. Jollyta et al., “Open Access Decision Tree and Chi-Square Analysis to Determine Student ’ s Criteria Choosing Study Program,” vol. 7, no. 10, pp. 12–20, 2018.
- [5] H. Hamsa et al, “Student Academic Performance Prediction Model Using Decision Tree and Fuzzy Genetic Algorithm,” *Procedia Technol.*, vol. 25, pp. 326–332, 2016.
- [6] D. Jollyta et al, “Analysis of Slow Moving Goods Classification Technique : Random Forest and Naïve Bayes,” vol. 5, no. 2, pp. 134–139, 2019.
- [7] N. H. Niloy et al, “Naïve Bayesian Classifier and Classification Trees for the Predictive Accuracy of Probability of Default Credit Card Clients,” vol. 3, no. 1, pp. 1–12, 2018.
- [8] B. Lowe et al, “M ULTISPECTRAL I MAGE A NALYSIS U SING R ANDOM,” vol. 6, no. 1, pp. 1–14, 2015.
- [9] W. E. Nurjanah et al, “Analisis Sentimen Terhadap Tayangan Televisi Berdasarkan Opini Masyarakat pada Media Sosial Twitter menggunakan Metode K-Nearest Neighbor dan Pembobotan Jumlah Retweet,” *J. Pengemb. Teknol. Inf. dan Ilmu Komput. Univ. Brawijaya*, vol. 1, no. 12, pp. 1750–1757, 2017.