

KLASIFIKASI TIPE GELANDANG SEPAK BOLA BERDASARKAN DATA KEMAMPUAN MENGGUNAKAN METODE NAIVE BAYES

Ni Wayan Wisswani¹⁾, *Tien Rahayu Tulili²⁾, Muhammad Farman Andrijasa³⁾, M. Rangga Satria N⁴⁾, Siti Jumani⁵⁾, Eko Wahyudi⁶⁾

¹⁾Program Studi Teknik Informatika, Politkenik Negeri Bali

^{2,4,6)}Program Studi Teknologi Rekayasa Komputer, Politeknik Negeri Samarinda

³⁾Program Studi Teknik Informatika Multimedia, Politeknik Negeri Samarinda

Email: wisswani@pnb.ac.id¹⁾, tien.tulili@polnes.ac.id²⁾, andrijasa@polnes.ac.id³⁾, mrangga2019@gmail.com⁴⁾, sitijumani898@gmail.com⁵⁾, ekobkrt@gmail.com⁶⁾

Abstrak - Sepak bola merupakan olahraga paling populer dan paling digemari di seluruh dunia. berdasarkan hasil survei yang dilakukan oleh Fédération Internationale de Football Association (FIFA) pada tahun 2001 menyatakan bahwa sepakbola adalah olahraga paling populer dimainkan hari ini. Survei ini menunjukkan bahwa lebih dari 240 juta orang memainkan olahraga sepak bola di lebih dari 200 negara di hampir setiap bagian dari dunia. Salah satu posisi dalam sepak bola adalah gelandang pemain tengah atau dalam Bahasa Inggris disebut dengan *midfielder* Posisi gelandang sepak bola terdiri dari beberapa tipe yaitu CMF (*central midfielder*), AMF (*attacking midfielder*), DMF (*defensive midfielder*), RMF (*right midfielder*), LMF (*left midfielder*). Namun, dalam penelitian ini, hanya tiga tipe yang akan digunakan yaitu CMF, DMF, dan AMF. Pada penelitian ini akan diangkat permasalahan yaitu: bagaimana seorang pemain gelandang dapat diklasifikasi kedalam posisi atau tipe gelandang yang tepat? Pendekatan yang digunakan yaitu menggunakan pendekatan klasifikasi menggunakan algoritma Naïve Bayes bertujuan untuk melakukan klasifikasi tipe gelandang sepak bola. Dari empat skenario yang dilaksanakan didapatkan bahwa hasil akurasi pada masing-masing skenario adalah sebesar 80%, 80%, 82.5%, dan 80.182%.

Kata Kunci – *klasifikasi, metode naive bayes, machine learning, kecerdasan buatan, weka*

1. PENDAHULUAN

Sepak bola merupakan olahraga paling populer dan paling digemari di seluruh dunia, pernyataan tersebut barangkali tidak terbantahkan, bahkan ada yang berpendapat bahwa tidak diperlukan sebuah penelitian ilmiah untuk mendapatkan pengesahan atas pernyataan tersebut. Berdasarkan hasil survei yang dilakukan oleh Fédération Internationale de Football Association (FIFA) pada tahun 2001 menyatakan bahwa sepakbola adalah olahraga paling populer dimainkan hari ini. Survei ini menunjukkan bahwa lebih dari 240 juta orang memainkan olahraga sepak bola di lebih dari 200 negara di hampir setiap bagian dari dunia. Berdasarkan survei diatas bisa dikatakan bahwa sepak bola merupakan olahraga yang paling digemari dan populer di seluruh dunia.

Salah satu posisi dalam sepak bola adalah gelandang, pemain tengah atau dalam Bahasa Inggris disebut dengan *midfielder*. *Midfielder* adalah pemain yang diposisikan di antara para bek dan penyerang tugas seorang gelandang adalah menghubungkan antara bek dan penyerang, yaitu mencegah agar serangan lawan tidak sampai ke pertahanan dan

menggiring bola sampai ke depan untuk diberikan kepada penyerang ketika tim dalam mode penyerangan. Para pemain di posisi ini biasanya paling banyak menguras tenaga. Karena jarak tempuh mereka dalam permainan yang cukup jauh. Hal ini karena gelandang adalah posisi yang menghubungkan antara lini belakang (*defenders*) dan lini depan (*strikers*).

Tetapi bagi orang awam *midfielder* hanya mempunyai satu posisi, semua pemain yang berada di lapangan tengah adalah seorang gelandang yang mempunyai tugas yang sama yaitu hanya menjalankan tugas seperti mengoper bola dan membantu penyerangan. Namun demikian, dalam menjalankan tugasnya posisi gelandang sendiri terbagi menjadi beberapa tipe, ada yang lebih condong ke bertahan dan ada juga yang kedudukannya hampir seperti pemain penyerang. *Midfielder* sebenarnya terbagi menjadi beberapa tipe antara lain: CMF (*central midfielder*), AMF (*attacking midfielder*), DMF (*defensive midfielder*), RMF (*right midfielder*), LMF (*left midfielder*).

Pengelompokkan pemain bola ke dalam posisi yang tepat tentu saja dilakukan oleh pelatih tim bola.

Pelatih ini biasanya akan menempatkan setiap pemain pada posisi yang tepat berdasarkan dari kemampuan setiap pemain bola, antara lain *dribbling*, *ball control*, *positioning*, *interception*, *sliding tackle*, *vision*, *passing*, dan stamina. Penempatan para pemain bola khususnya ke dalam posisi CMF, AMF, DMF, RMF, maupun LMF dapat juga dilakukan dengan menggunakan pendekatan menggunakan teknik pembelajaran mesin. Salah satu teknik pembelajaran mesin yang dapat digunakan yaitu menggunakan pendekatan klasifikasi. Adapun dalam penelitian ini, metode klasifikasi yang digunakan adalah metode Naïve Bayes.

Beberapa penelitian yang telah dilakukan sebelumnya antara lain oleh [1] yaitu melakukan klasifikasi berdasarkan data penduduk miskin yang diperoleh dari Kecamatan Tibawa dengan menggunakan teknik data mining. Atribut yang akan digunakan dalam melakukan klasifikasi penduduk adalah Umur, Pendidikan, Pekerjaan, Penghasilan, Tanggungan, Status (Kawin/Belum Kawin). Metode yang digunakan adalah metode *Naïve Bayes Classifier*, yang merupakan salah satu teknik pengklasifikasian dalam *data mining*. Berdasarkan penelitian yang dilakukan dihasilkan kesimpulan bahwa, Sistem klasifikasi masyarakat miskin di wilayah pemerintahan Kecamatan Tibawa Kab. Gorontalo dapat direkayasa dan Berdasarkan hasil pengujian *confusion matrix* dengan teknik split validasi, penggunaan metode klasifikasi *naïve bayes* terhadap dataset yang telah diambil pada objek penelitian diperoleh tingkat akurasi sebesar 73% atau termasuk dalam kategori *Good*. Sementara nilai *Precision* sebesar 92% dan *Recall* sebesar 86%.

Penelitian yang dilakukan oleh [2][3] yaitu tentang klasifikasi nasabah dengan menggunakan metode Naïve Bayes yaitu pendapatan untuk perusahaan asuransi ditentukan oleh jumlah premi yang dibayar oleh nasabah. Banyaknya nasabah yang tidak lancar membayar premi berpengaruh terhadap kinerja serta eksistensi perusahaan sehari-hari. Algoritma Naive Bayes berbasis Backward Elimination bertujuan untuk melakukan klasifikasi nasabah asuransi dengan hasil akurasi 85,89 % dengan delapan atribut weight yaitu umur, jangka waktu, cara bayar, premi, jumlah hari, pekerjaan, penghasilan dan mata uang.

Penelitian yang dilakukan oleh [4] yaitu tentang penentuan konsentrasi siswa dilakukan agar setiap siswa dapat belajar lebih banyak dalam mata pelajaran sesuai dengan konsentrasi yang telah ditentukan untuk masing-masing mahasiswa. Untuk menentukan konsentrasi siswa-siswa ini, agar dapat diterapkan metode Naive Bayes dalam mengklasifikasikan konsentrasi siswa berdasarkan data yang dilatih untuk menentukan probabilitas setiap kriteria yang digunakan dalam kriteria nominal serta kriteria numerik dan dari nilai probabilitas dapat diprediksi konsentrasi yang sesuai untuk siswa. Berdasarkan hasil penelitian yang dilakukan untuk mengetahui konsentrasi siswa dengan metode Naive Bayes, keakuratan hasil klasifikasi yang diperoleh sesuai dengan data yang diujikan sebesar 90,8333%, dimana data 120 siswa yang diuji, terdapat sebanyak 109 siswa data berhasil diklasifikasikan dan diprediksi dengan benar

Penelitian yang dilakukan oleh [5] yang berjudul “penerapan metode naïve bayes dalam klasifikasi kelayakan keluarga penerima beras rastra” Beras rastra adalah sebuah program pemerintah yang bertujuan untuk meringankan beban keluarga miskin atau hampir miskin dalam hal pangan. Dalam implementasinya pada desa Bandar Siantar Kecamatan Gunung Malela masih belum optimal dikarenakan masih banyak program rastra yang belum tepat sasaran. Konsep data mining akan mempermudah mengatasi masalah yang belum optimal di desa Bandar Siantar Kecamatan Gunung Malela. Maka, metode klasifikasi mampu menemukan model yang membedakan konsep atau kelas data, dengan tujuan untuk dapat memperkirakan kelas dari suatu objek yang labelnya tidak diketahui. Oleh sebab itu algoritma Naive Bayes dapat memprediksi peluang di masa depan berdasarkan pengalaman dimasa sebelumnya, pada penelitian ini peneliti mengambil data latih sebanyak 70 data dan sebuah data uji, dengan menggunakan 6 kriteria yaitu Status PKH, Jumlah Tanggungan, Kepala rumah Tangga, Kondisi Rumah, Jumlah Penghasilan, dan Status Pemilik Rumah. Hasil penelitian ini diharapkan dapat membantu pemerintah khususnya di daerah dalam menentukan kelayakan keluarga penerima beras Rastra.

Dan juga beberapa penelitian lainnya yang mengimplementasikan metode Naïve Bayes untuk mengklasifikasi data pada beberapa area kehidupan

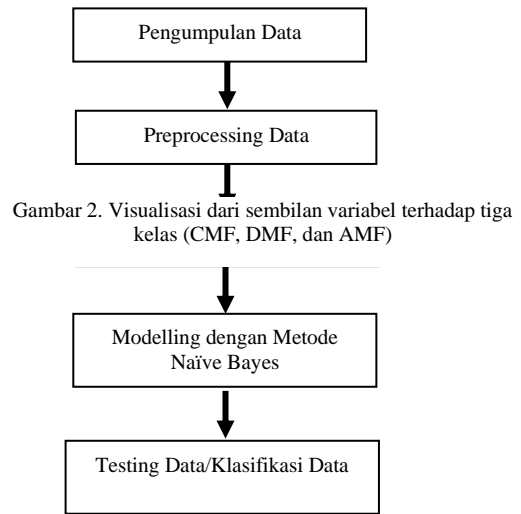
disekitar kita membuktikan bahwa saat ini penambahan data dengan metode pendekatan pembelajaran mesin menjadi cukup populer diimplementasikan [6][7][8][9][10].

Pada penelitian ini, penempatan pemain bola khususnya pemain dengan posisi gelandang tengah akan diklasifikasi ke dalam tiga posisi gelandang tengah yaitu AMF, DMF, dan CMF dengan menggunakan metode Naïve Bayes.

2. METODOLOGI PENELITIAN

Penelitian ini dilaksanakn sesuai dengan metodologi penelitian yang digambarkan pada gambar 1. Pada pengumpulan data, data yang digunakan yaitu Data kemampuan seorang gelandang sepak bola berdasarkan game FIFA dan didapat dari website resmi EA SPORTS [10], yaitu data pemain sepak bola yang berposisi sebagai gelandang yang masih aktif bermain. Jumlah data gelandang sepak bola yang digunakan berjumlah 101 data. Data inilah yang disebut sebagai dataset. Tipe gelandang sepak bola yang digunakan dalam penelitian ini adalah CMF, AMF, DMF.

Selanjutnya, data akan dibersihkan dan disesuaikan dengan format data yang dapat diproses oleh tool Weka. Kemudian, pada tahap training data, dataset akan dibagi menjadi beberapa bagian. Pada penelitian ini, prosentasi pembagian menjadi data latih dan data uji masing-masing sebesar 70% dan 30%; 60% dan 40%; 50% dan 50%. Metode pembagian data latih dan data uji dilakukan secara acak. Setelah itu, tool Weka akan membangun model berdasarkan data latih yang ada dengan menggunakan metode Naïve Bayes. Demikian juga halnya, saat melakukan klasifikasi terhadap data uji, metode yang digunakan adalah metode Naïve Bayes.



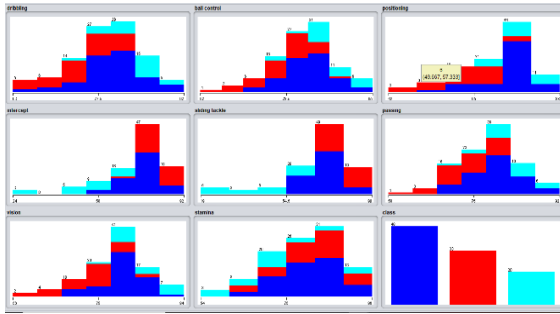
Gambar 2. Visualisasi dari sembilan variabel terhadap tiga kelas (CMF, DMF, dan AMF)

Gambar 1 Kerangka Konsep Penelitian

3. HASIL DAN PEMBAHASAN

Pengumpulan data yang dilakukan pada penelitian ini berupa data awal gelandang sepak bola yang telah dikumpulkan dari website resmi EA SPORTS. Data yang digunakan hanyalah gelandang sepak bola yang telah dibagi menjadi 3 tipe yaitu : Central Midfield (CMF), Attacking Midfield (AMF), dan Defensive Midfield (DMF). Variabel data yang digunakan sebanyak Sembilan(9) variabel antara lain: *dribbling*, *ball control*, *positioning*, *interception*, *sliding tackle*, *vision*, *passing*, *stamina*, dan *class*. Banyaknya data yang digunakan adalah 110 pemain. Tool yang digunakan pada proses awal sampai dengan proses akhir adalah tool *Weka* versi 3.8.3. Dataset yang digunakan disimpan dalam bentuk berkas berekstensi *.csv*.

Hasil preprocessing dapat dilihat pada gambar 2 dibawah. Pada gambar 2, visualisasi dari masing-masing variable menunjukkan banyaknya jumlah data yang dikelompokkan berdasarkan tiga kelas yaitu: CMF, AMF, dan DMF. CMF ditunjukkan dengan grafik batang berwarna biru, DMF ditunjukkan dengan grafik batang berwarna merah, dan AMF ditunjukkan dengan grafik batang berwarna magenta.



Pada pelatihan, pemodelan, dan pengklasifikasian data uji menggunakan 4 macam skenario, yaitu tiga skenario pertama menggunakan metode *splitting data* yaitu membagi dataset kedalam data latih dan data uji dengan menggunakan 3 jenis prosentase berbeda dengan rasio data latih/data uji yaitu 80/20, 70/30, dan 60/40. Namun, satu skenario yang terakhir menggunakan metode *cross validation 10-folds*.

Hasil klasifikasi masing-masing skenario dapat dilihat pada tabel 1-3 berikut.

Tabel 1. Hasil Klasifikasi Data dengan skenario 1

| Skenario 1 | Precision (%) | Recall (%) | F-Measure (%) |
|------------|---------------|------------|---------------|
| Kelas CMF | 60 | 100 | 75 |
| Kelas DMF | 100 | 60 | 75 |
| Kelas AMF | 100 | 100 | 100 |

Pada skenario 1, akurasi ketepatan klasifikasi yang didapatkan sebesar 80% dengan hasil Precision, Recall dan F-Measure dapat dilihat pada table 1 diatas. Berdasarkan table 1 yaitu pada kelas AMF memiliki prosentase F-Measure paling tinggi yaitu sebesar 100%.

Tabel 2. Hasil Klasifikasi Data dengan skenario 2

| Skenario 1 | Precision (%) | Recall (%) | F-Measure (%) |
|------------|---------------|------------|---------------|
| Kelas CMF | 68.8 | 9.17 | 78.6 |
| Kelas DMF | 100 | 64.3 | 78.3 |
| Kelas AMF | 80 | 100 | 88.9 |

| Skenario 1 | Precision (%) | Recall (%) | F-Measure (%) |
|------------|---------------|------------|---------------|
| Kelas CMF | 68.8 | 9.17 | 78.6 |
| Kelas DMF | 100 | 64.3 | 78.3 |
| Kelas AMF | 80 | 100 | 88.9 |

Pada skenario 70/30, hasil akurasi klasifikasi yang didapatkan sebesar 80%. Tetapi nilai F-measure yang didapatkan tiap kelas berbeda dengan hasil yang didapatkan pada skenario 1, dimana untuk nilai F-Measure pada kelas CMF dan DMF masing-masing meningkat menjadi 78.6% dan 78.3%. Sedangkan untuk hasil F-Measure dari kelas AMF, mengalami penurunan sebesar sekitar 10% lebih.

Table 3. Hasil Klasifikasi Data dengan skenario 3

| Skenario 1 | Precision (%) | Recall (%) | F-Measure (%) |
|------------|---------------|------------|---------------|
| Kelas CMF | 80 | 75 | 77.4 |
| Kelas DMF | 81.3 | 81.3 | 81.3 |
| Kelas AMF | 88.9 | 100 | 94.1 |

Pada skenario 60/40, hasil akurasi klasifikasi yang didapatkan naik menjadi 82.5%. Tetapi nilai F-measure yang didapatkan tiap kelas berbeda dengan hasil yang didapatkan pada skenario 1, dimana untuk nilai F-Measure pada kelas CMF turun menjadi 77.4%, DMF meningkat menjadi 81.3%. Sedangkan untuk hasil F-Measure dari kelas AMF, mengalami peningkatan menjdai 94.1%.

Table 4. Hasil Klasifikasi Data dengan skenario 4

| Skenario 1 | Precision (%) | Recall (%) | F-Measure (%) |
|------------|---------------|------------|---------------|
| Kelas CMF | 78.4 | 83.3 | 80.8 |
| Kelas DMF | 77.4 | 72.7 | 75.0 |
| Kelas AMF | 89.5 | 85.0 | 87.2 |

Pada skenario terakhir, yaitu menggunakan *cross-fold validation*, hasil akurasi yang didapatkan sebesar 80.182%. Jika dibandingkan dengan hasil skenario 4, selisih yang didapatkan sebesar kurang lebih 1.3% lebih rendah.

4. Kesimpulan

Dari semua percobaan yang dilakukan sebelumnya, didapatkan bahwa prosentase pembagian data latih dan data uji dapat mempengaruhi hasil akurasi klasifikasi yaitu semakin besar prosentasi data latih, semakin kecil prosentasi akurasi yang didapatkan. Namun, saat *cross-validation* yang

digunakan yaitu dengan nilai k sebesar 10, hasil akurasi yang didapatkan kurang lebih dengan metode splitting dengan rasio 60/40. Hasil akurasi dari scenario 1, 2, 3, dan 4 didapatkan masing-masing sebesar 80%, 80%, 82.5%, dan 80.182%.

Namun demikian, percobaan ini hanya dilaksanakan pada empat scenario dan untuk mendapatkan hasil perbandingan yang lebih baik, diharapkan dapat dilakukan percobaan yang lebih banyak dan menambah jumlah instans pada dataset lebih banyak lagi sehingga akan didapatkan model yang lebih baik.

REFERENSI

- [1] H. Annur, "Klasifikasi Masyarakat Miskin Menggunakan Metode Naive Bayes," *Ilk. J. Ilm.*, 2018, Doi: 10.33096/Ilkom.V10i2.303.160-165.
- [2] B. Betrisandi, "Klasifikasi Nasabah Asuransi Jiwa Menggunakan Algoritma Naive Bayes Berbasis Backward Elimination," *Ilk. J. Ilm.*, 2017, Doi: 10.33096/Ilkom.V9i1.116.96-101.
- [3] "Penerapan Algoritma Naive Bayes Untuk Mengklasifikasi Data Nasabah Asuransi," 2014, Doi: 10.26555/Jifo.V8i1.A2086.
- [4] A. Saleh, "Klasifikasi Metode Naive Bayes Dalam Data Mining Untuk Menentukan Konsentrasi Siswa (Studi Kasus Di Mas Pab 2 Medan)," *Konf. Nas. Pengemb. Teknol. Inf. Dan Komun. 2014*, 2014, Doi: 10.1007/Bf03053466.
- [5] C. Fadlan, S. Ningsih, And A. P. Windarto, "Penerapan Metode Naive Bayes Dalam Klasifikasi Kelayakan Keluarga Penerima Beras Rastra," *J. Tek. Inform. Musirawas*, 2018, Doi: 10.32767/Jutim.V3i1.286.
- [6] A. Saleh, "Implementasi Metode Klasifikasi Naive Bayes Dalam Memprediksi Besarnya Penggunaan Listrik Rumah Tangga," *Creat. Inf. Technol. J.*, 2015.
- [7] T. I. Andini, W. Witanti, And F. Renaldi, "Prediksi Potensi Pemasaran Produk Baru Dengan Metode Naive Bayes Classifier Dan Regresi Linear," *Semin. Nas. Apl. Teknol. Inf.*, 2016.
- [8] A. Saleh, "Penerapan Data Mining Dengan Metode Klasifikasi Naive Bayes Untuk Memprediksi Kelulusan Mahasiswa Dalam Mengikuti English Proficiency Test (Studi Kasus : Universitas Potensi Utama)," *Konf. Nas. Sist. Informasi, Univ. Klabat, Manado, Indonesia, Vol. 2015*, 2015.
- [9] S. Widaningsih, "Perbandingan Metode Data Mining Untuk Prediksi Nilai Dan Waktu Kelulusan Mahasiswa Prodi Teknik Informatika Dengan Algoritma C4,5, Naive Bayes, Knn Dan Svm," *J. Tekno Insentif*, 2019, Doi: 10.36787/Jti.V13i1.78.

- [10] M. S. Mustafa, M. R. Ramadhan, And A. P. Thenata, "Implementasi Data Mining Untuk Evaluasi Kinerja Akademik Mahasiswa Menggunakan Algoritma Naive Bayes Classifier," *Creat. Inf. Technol. J.*, 2018, Doi: 10.24076/Citec.2017v4i2.106.
- [11] <https://www.easports.com/fifa/ultimate-team/fut/database> diakses tanggal 15 Desember 2019