

PENENTUAN VALIDASI DATA PEMILIH DAN KLASIFIKASI HASIL PEMILU DPRD KAB.BONE UNTUK MEMPREDIKSI PARTAI PEMENANG MENGGUNAKAN METODE NAIVE BAYES

Woro Isti Rahayu¹, Aulyardha Anindita², Mohamad Nurkamal Fauzan³

Program Studi D4 Teknik Informatika¹²³
Politeknik Pos Indonesia¹²³

email: ¹⁾ woroisti@poltekpos.ac.id , ²⁾ aulyardhaanindita@gmail.com, ³⁾ m.nurkamal.f@poltekpos.ac.id

Abstrak

Validasi data pemilih merupakan salah satu kegiatan yang wajib dilakukan oleh Komisi Pemilihan Umum yang dilakukan setiap bulannya. KPU Kab. Bone memiliki kendala yang didapatkan yaitu proses pengiriman data tidak cepat dan memiliki data yang double. Maka dari itu, dibutuhkan suatu sistem yang bisa menjadi wadah guna mengatasi masalah tersebut. Selain itu, dalam pemilu tentunya tidak lepas dari partai politik. beberapa partai penguasa yang selalu menjadi pemenang dalam setiap pemilu membutuhkan adanya prediksi partai pemenang pada pemilihan selanjutnya, begitu juga dengan partai yang kalah tentunya ingin meningkatkan kredibilitasnya dan menjadi pemenang pada pemilihan selanjutnya. sehingga dari hasil penelitian ini dapat diketahui bahwa dengan menggunakan metode naive bayes dapat memprediksi partai pemenang di Kab.Bone dengan akurasi sebesar 93%.

Kata Kunci: Validasi, Pemilih, Prediksi, Partai, Naive Bayes

Abstract

Validation of voter data is one of the activities that must be carried out by the General Elections Commission which is carried out every month. District Commission. Bone has a problem that is obtained, namely the process of sending data is not fast and has double data. Therefore, we need a system that can be a forum to overcome these problems. In addition, in the election of course can not be separated from political parties. some ruling parties that always win in every election need a prediction of the winning party in the next election, as well as the losing party of course want to increase their credibility and become the winner in the next election. so from the results of this study it can be seen that using the naive bayes method can predict the winning party in Bone Regency with an accuracy of 93%.

Keywords: Validation, Voter, Prediction, Party, Naive Bayes

1. PENDAHULUAN

Pemilihan umum (PEMILU) merupakan suatu peristiwa politik yang sangat penting dalam menentukan pemimpin dalam negara demokrasi[1]. Terdapat program yang dilaksanakan dalam optimalisasi pelayanan publik pada Pemilu kepala daerah yaitu memfasilitasi pemilih dalam penggunaan hak pilih. Agar seseorang masuk dalam daftar pemilih, masyarakat harus terdata dulu di dinas kependudukan dan catatan sipil, dan jika telah terdata selanjutnya penyelenggara pemilu akan

memutakhirkan data yang bersangkutan apakah dia memiliki hak pilih atau bisa menggunakan hak pilihnya atau mempunyai hak pilih tapi tidak diperbolehkan menggunakan hak pilihnya karena menjadi TNI/POLRI atau mempunyai hak pilih tetapi harus berpindah tempat pencoblosannya karena masalah pekerjaan atau kegiatan lainnya[2].

Didalam Pemilu sendiri, tentunya tidak lepas dari Partai Politik. Berdasarkan kutipan dari Miriam Budiardjo, Partai Politik merupakan suatu sarana bagi warga negara

untuk turut serta atau berpartisipasi dalam proses pengelolaan negara. Dan partisipasi warga negara ini jika disampaikan perorangan terkadang tidak didengar bahkan hilang tak berbekas, maka dari itu jika aspirasi bersama ditampung dan diolah melalui partai politik sehingga menjadi kebijakan partai dan tentunya akan menjadi program kerja, yang disampaikan wakil rakyatnya kepada pemerintah untuk menjadi kebijakan umum mengatasi berbagai permasalahan di masyarakat. [3]

Validasi data pemilih di KPU Kab.Bone masih tetap dilakukan tiap bulannya, tapi tentunya terdapat beberapa kendala yang sering didapatkan dalam melakukan proses validasi data pemilih. Diantaranya, berdasarkan hasil wawancara didapatkan bahwa hampir sekitar 10% ditemukan data yang double pada proses pemutakhiran data pemilih. Hal itu disebabkan karena pindah domisili, meninggal, memiliki ktp yang double dan lain sebagainya. Selain itu, beberapa data yang double tersebut terjadi karena pelaporan dari tiap kecamatan tidak melakukan pengecekan terlebih dahulu dan langsung melaporkan data tersebut ke capil, sehingga data di capil banyak yang double karena sistemnya juga masih dilakukan secara manual, dan data tersebut yang diambil oleh pegawai KPU untuk melakukan validasi data pemilih tiap bulannya.

Sehingga untuk mengatasi hal tersebut diterapkan metode triangulasi, metode triangulasi adalah salah satu metode yang bisa digunakan untuk menguji suatu informasi yang dikatakan valid ataupun tidak terhadap informasi yang diperoleh dari beberapa riset yang telah dilakukan. Dan dari metode ini didapatkan bahwa untuk mengecek validasi data pemilih digunakan NIK, NIK digunakan karena tiap masyarakat memiliki NIK yang berbeda-beda. NIK digunakan agar setiap data laporan yang dilaporkan oleh pihak kecamatan sudah dicek lebih awal sehingga permasalahan double data dapat diatasi.

Partai politik memiliki peran penting dalam pemilu. Beberapa partai berlomba untuk menjadi partai penguasa. Berbagai macam usaha yang dilakukan oleh partai-partai tersebut untuk mempertahankan atau untuk meningkatkan kredibilitasnya dalam masyarakat, sehingga pastinya mereka membutuhkan suatu prediksi yang bisa mengetahui bahwa misalnya di daerah A mereka akan menang atau kalah. Sehingga

lebih awal mereka bisa menyusun suatu rencana atau strategi yang bisa mereka lakukan untuk mewujudkan tujuan mereka.

Permasalahan tersebut bisa diselesaikan dengan data mining. Data mining merupakan salah satu cara yang bisa digunakan untuk memprediksi dan mendeteksi suatu kasus, termasuk memprediksi partai pemenang. Terdapat berbagai macam algoritma atau metode dalam data mining yang bisa digunakan dalam memprediksi diantaranya, metode regresi linear, decision tree, algoritma k-means, dan lain sebagainya. [4]

Pada penelitian sebelumnya yang dilakukan oleh [5] melakukan perbandingan metode data mining yaitu algoritma C4.5, Support vector machine (SVM), K-nearest neighbor (kNN) dan Naive bayes untuk memprediksi tingkat kelulusan prodi Teknik Informatika dengan standar yang telah ditetapkan di Universitas Suryakencana. Hasil akhir dari keempat algoritma tersebut diperoleh bahwa metode algoritma Naive bayes merupakan salah satu algoritma terbaik untuk memprediksi kelulusan mahasiswa yang tepat waktu dengan nilai accuracy (76,79%), error (23,17%), dan AUC (0,850).

Sedangkan pada penelitian yang dilakukan oleh [6], melakukan perbandingan kinerja metode naive bayes dan kNN untuk mengklasifikasi dokumen yang dapat mengelompokkan artikel secara otomatis dan akurat. Dari hasil penelitian tersebut menunjukkan bahwa metode naive bayes memiliki kinerja yang lebih baik dibandingkan metode kNN dengan tingkat akurasi 70% sedangkan untuk metode kNN memiliki tingkat akurasi yang cukup rendah yaitu 40%.

Maka dari itu, untuk mengatasi kedua masalah tersebut dibuatlah sebuah sistem atau aplikasi untuk mengecek validasi data pemilih dan prediksi kemenangan partai di Kab.Bone. Dengan melakukan pengecekan data pemilih maka data pemilihnya akan valid sehingga hasil prediksinya juga valid karena sesuai dengan data pemilih hal itu disebabkan karena data pemilih sangat mempengaruhi hasil prediksi dari tiap partai. Untuk mengecek validasi data pemilih digunakan metode triangulasi sedangkan untuk memprediksi partai pemenang digunakan metode naive bayes.

Dari beberapa penelitian sebelumnya, Metode naive bayes memiliki tingkat akurasi lebih tinggi dibandingkan dengan metode-

metode lain dalam melakukan prediksi. Metode Naive bayes merupakan salah satu metode yang bersifat sederhana dan efektif dalam pengolahan data yang besar. Naive bayes adalah salah satu metode klasifikasi yang berakar pada Teorema Bayes. Teorema Bayes ini dikombinasikan dengan Naive yang berarti setiap atribut atau variabel nya bersifat bebas. Metode ini memiliki akurasi dan kecepatan yang tinggi jika diaplikasikan pada data yang besar. Pada metode ini dilakukan dengan memanfaatkan data training untuk menghasilkan probabilitas setiap kriteria untuk kelas yang berbeda dan selanjutnya nilai-nilai probabilitas dari kriteria tersebut akan dioptimalkan untuk memprediksi partai pemenang di Kab.Bone.

Dengan adanya penelitian penelitian ini diharapkan nantinya akan membantu pihak-pihak yang membutuhkan prediksi partai pemenang khususnya pihak partai politik itu sendiri dalam mempertahankan ataupun meningkatkan kredibilitasnya sehingga mereka dapat merencanakan lebih awal tentang strategi-strategi yang bisa dilakukan dalam meningkatkan atau mempertahankan suaranya. Serta dengan adanya penelitian penelitian ini juga dapat membantu pegawai KPU dalam melakukan validasi data pemilih tiap bulannya dengan mempercepat proses validasi data pemilih dan mengatasi permasalahan double data atau kesalahan input data.

2. LANDASAN TEORI

A. Pemilih

Berdasarkan Undang-Undang Nomor 8 Tahun 2015 yang dijelaskan pada pasal 1 ayat (6) menjelaskan bahwa Pemilih merupakan penduduk yang berusia paling rendah 17 tahun atau sudah atau pernah kawin yang terdaftar dalam pemilihan. Kemudian pada pasal 57 ayat (1) ditegaskan bahwa untuk dapat menggunakan hak pilih, warga negara Indonesia harus terdaftar sebagai pemilih dan pada ayat (3) dijelaskan bahwa untuk didaftar sebagai pemilih, warga negara Indonesia harus memenuhi syarat yaitu, tidak sedang terganggu jiwa/ingatannya, tidak sedang dicabut hak pilihnya berdasarkan putusan pengadilan yang mempunyai kekuatan hukum tetap [7]

B. Klasifikasi

Klasifikasi merupakan suatu proses dalam menemukan model atau suatu fungsi yang dapat menjelaskan atau membedakan konsep atau kelas data yang memiliki tujuan untuk memperkirakan kelas yang tidak diketahui dari suatu objek. [8]

C. Pemilu

Pemilihan Umum (Pemilu) adalah salah satu instrumen penting dalam suatu negara demokrasi yang menganut sistem perwakilan yang berfungsi sebagai alat penyaring politikus yang akan mewakili dan membawa suara rakyat didalam lembaga perwakilan [9]

D. Prediksi

Prediksi merupakan suatu kegiatan yang dilakukan untuk memprediksi atau memperkirakan segala sesuatu yang akan terjadi kedepannya, prediksi ini digunakan sebagai informasi, ulasan dan pengetahuan berdasarkan faktor-faktor yang mempengaruhi berdasarkan fakta-fakta yang ada. [10]

E. Partai Politik

Partai berasal dari dan kata part yang berarti bagian dan menunjukkan kepada bagian dari para warga negara. Sedangkan kata partai sendiri menunjukkan sekumpulan orang-orang. Jadi, bisa disimpulkan partai yaitu menunjukkan kepada sekumpulan sejumlah warga dan suatu negara yang menggabungkan satu kesatuan yang mempunyai tujuan tertentu. [11]

F. DPRD

DPRD merupakan singkatan dari Dewan Perwakilan Rakyat Daerah. DPRD merupakan suatu lembaga perwakilan tempat masyarakat mengeluarkan atau menyampaikan aspirasi dan menyuarakan kepentingannya.[13]

G. Data

Menurut Webster's New World's Dictionary, datum yaitu something known or assumed, yang artinya datum (dalam bentuk tunggal) adalah sesuatu yang diketahui atau dianggap. Sehingga dengan demikian, data dapat memberi gambaran tentang suatu keadaan atau persoalan. [12]

H. Akurasi

Akurasi merupakan nilai yang diambil dari seberapa dekat nilai hasil pengukuran dengan nilai sebenarnya atau true value

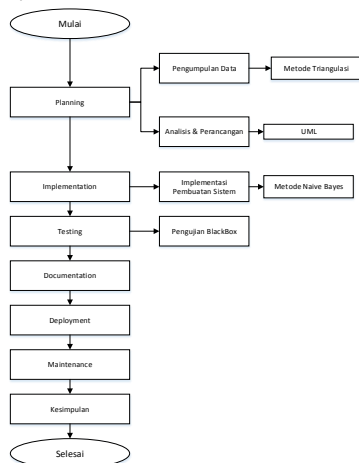
atau nilai yang biasa dianggap benar (accepted value). [14]

I. Database

Basis data dapat juga diartikan sebagai kumpulan data yang berisi tentang suatu benda/kejadian yang saling berhubungan antar satu sama lain. [15]

3. METODE PENELITIAN

Dalam perancangan suatu sistem, pada penelitian ini dibuat diagram alur yang menjelaskan secara singkat tahap-tahap alur penelitian yang akan dilakukan. Diagram alur metodologi penelitian ini mengikuti metode pengembangan sistem yang digunakan yaitu Metode Agile Software Development. Berikut merupakan diagram alur metodologi penelitian :



Gambar 1. Diagram Alur Metodologi Penelitian

Adapun tahapan-tahapan metodologi penelitian yang dilakukan yaitu :

1. Planning, Pada tahapan ini, dalam membangun sebuah sistem terlebih dahulu kita harus melakukan perencanaan diantaranya dengan melakukan pengumpulan data-data yang mendukung dalam pembuatan sistem yang akan dibangun. Pengumpulan data disini menggunakan metode triangulasi dalam mengumpulkan data yang akan diolah. Setelah melakukan pengumpulan data, data yang didapatkan tersebut akan dicek validasi datanya dengan menggunakan metode triangulasi. Setelah itu, dilakukan proses perancangan sistem untuk menjawab kebutuhan pengguna melalui flowmap/flowchart dan diagram UML.
2. Implementation, Pada tahap ini, rancangan design yang telah dilakukan pada tahap

sebelumnya akan diubah menjadi kode-kode program yang nantinya akan diintegrasikan menjadi sebuah sistem. Dan pada pembuatan sistem ini menggunakan bahasa pemrograman PHP serta menggunakan metode Naive bayes untuk memprediksi partai pemenang di Kab.Bone Sulawesi Selatan.

3. Testing, Pada tahap ini, sistem yang telah dibuat akan diintegrasikan dan dites untuk menguji apakah sistem tersebut dapat berfungsi dengan baik. Pada pengujian sistem ini, menggunakan metode black box yaitu metode pengujian yang dilakukan pada tampilan program apakah program dapat berjalan dengan baik dan sesuai dengan apa yang diinginkan, dengan kata lain yang dilihat hanyalah berdasarkan keluaran yang dihasilkan dari data atau kondisi masukan yang diberikan untuk fungsi yang ada tanpa melihat bagaimana prosesnya.
4. Documentation, Pada tahap ini, dilakukan dokumentasi. Dokumentasi ini sangat penting untuk dilakukan dalam pengembangan sistem. Yaitu dilakukan dokumentasi mengenai modul dan fungsi-fungsi yang ada pada sistem yang telah dibangun, hal ini dilakukan sebagai catatan pada saat pengembangan dan hal ini juga dilakukan untuk mempermudah tim dalam melakukan proses pengembangan selanjutnya. Proses yang dilakukan dalam tahap ini yaitu membuat dokumentasi modul dan fungsi-fungsi program, membackup file-file yang akan digunakan pada saat pengembangan, rancangan database, subrutim, nilai input dan juga nilai output.
5. Deployment, Pada tahap ini, dilakukan dengan menyediakan sistem yang telah berhasil dibuat untuk digunakan oleh end-user yaitu pegawai KPU Kab.Bone, Pegawai Kecamatan, dan Pegawai CAPIL
6. Maintenance, Pada tahap ini, sistem yang berhasil dibuat telah dipasang dan digunakan oleh end-user sehingga pada proses ini, akan dilakukan proses Maintenance atau pemeliharannya sendiri dengan memeriksa kembali kesalahan-kesalahan atau error yang tidak didapatkan pada tahap-tahap sebelumnya, sehingga ketika sistem tersebut telah digunakan secara langsung oleh end-user dapat atau

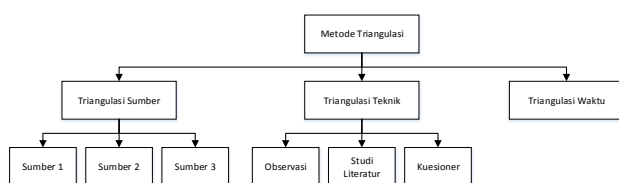
mampu memperbaiki kesalahan-kesalahan yang didapatkan pada proses ini, serta jika terdapat penambahan fitur-fitur baru akan dilakukan proses Maintenance atau pemeliharaan kembali.

7. Kesimpulan, Pada tahap kesimpulan dan saran merupakan tahapan terakhir dari kegiatan penelitian ini. Dalam tahapan kesimpulan sendiri dijelaskan secara singkat mengenai hasil dari penelitian yang telah dilakukan, sedangkan dalam tahap saran sendiri merupakan tahapan yang memberikan catatan jika ada kekurangan atau hal yang tidak sempurna dalam sebuah penelitian yang dilakukan.

Adapun penjelasan dari metode yang digunakan adalah sebagai berikut :

1. Metode Triangulasi

Pada tahap pengumpulan data di penelitian ini menggunakan metode triangulasi. Metode triangulasi merupakan suatu metode pengumpulan data yang memiliki sifat menggabungkan dari berbagai teknik pengumpulan data dan sumber data yang telah ada. Proses triangulasi ini akan dilakukan secara terus menerus sepanjang proses mengumpulkan data dan analisis data, sampai sudah tidak menemukan ada lagi perbedaan-perbedaan dan tidak ada lagi yang perlu dikonfirmasi kepada informan. Adapun tahap-tahap yang dilakukan dalam metode triangulasi yaitu :



Gambar 2. Metode Triangulasi

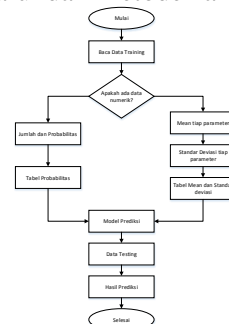
- a. Triangulasi Sumber, dilakukan untuk menguji kredibilitas data yang didapatkan dengan mengecek data yang telah diperoleh melalui beberapa sumber. Dari beberapa wawancara yang telah dilakukan dari ketiga sumber, dapat diperoleh kesimpulan bahwa terdapat beberapa kendala yang sering dialami oleh kpu kab.bone diantaranya seringkali ditemukan data yang double hampir sekitar 10% hal ini disebabkan karena data awal yang dilaporkan oleh pihak kecamatan tidak melakukan pengecekan terlebih dahulu, sehingga data double tersebut tidak dapat dihindari. Maka dari itu, dari beberapa wawancara

diusulkan NIK untuk mengatasi permasalahan double data tersebut, karena NIK tiap orang berbeda-beda maka laporan awal yang dilakukan oleh pihak kecamatan sudah dilakukan pengecekan dari awal sehingga ketika NIK yang diinputkan sama, maka otomatis laporan tersebut tidak akan dilaporkan sehingga dapat mengatasi permasalahan double data dan kesalahan dalam penginputan data.

- b. Triangulasi Teknik, Selain dengan wawancara, untuk memperkuat validitas datanya, dibutuhkan sumber lain, yaitu studi literatur yang mendukung akan data yang didapatkan. Disini penulis mendapatkan studi literatur dari jurnal dan buku-buku yang berkaitan dengan tema penelitian ini. Selain itu juga, penulis melakukan observasi, yaitu melakukan pengamatan secara langsung tentang objek yang diteliti dengan instansi terkait yaitu Kantor KPU Kab.Bone dan didampingi secara langsung oleh Pak Nasruddin Zaelany selaku anggota KPU Kab.Bone, untuk mengumpulkan segala informasi yang berkaitan dengan permasalahan-permasalahan yang ada, serta informasi yang dikumpulkan juga berkaitan dengan proses yang sedang berjalan saat ini.
- c. Triangulasi Waktu, Waktu juga sering mempengaruhi kredibilitas data. Misalnya data yang dikumpulkan dengan teknik wawancara di pagi hari pada saat narasumber masih segar, belum memiliki banyak pikiran atau masalah akan memberikan data yang lebih valid sehingga lebih kredibel. Maka dari itu, penulis melakukan wawancara kepada narasumber di pagi hari sekitar pukul 10.00 WITA

2. Metode Naive Bayes

Adapun alur dari metode naive bayes yaitu :



Gambar 3. Alur Metode Naive Bayes

a. Baca data training, Untuk menentukan suatu data yang nantinya akan dianalisis dengan metode naive bayes terlebih dahulu membaca data training atau data latihnya. Data training yang digunakan dalam penelitian ini adalah data hasil pemilihan DPRD Kab.Bone Tahun 2019, 2014, dan 2009

b. Mengecek type data, Setelah membaca data training, selanjutnya mengecek apakah data training tersebut memiliki data numerik atau hanya memiliki data categorical. Pada penelitian ini, data training yang digunakan memiliki data numerik yaitu pada atribut jumlah pemilih. Sedangkan untuk data training yang typenya categorical yaitu pada atribut nama partai, kecamatan, tahun, dan hasil.

c. Menghitung Nilai Mean dan Standar Deviasi, untuk menghitung nilai mean dan standar deviasi dari atribut yang type datanya numerik digunakan rumus :

$$\mu = \sum_i^n = 1x_i \text{ atau } \mu = \frac{x_1+x_2+x_3+\dots+x_n}{n}$$

Keterangan :

μ = rata-rata (mean)

i = nilai sampel ke-i

n = jumlah sampel

$$\sigma = \sqrt{\frac{\sum_{i=1}^n (x_i - \mu)^2}{n - 1}}$$

Keterangan :

σ = Standar deviasi

xi = nilai x ke-i

μ = rata-rata hitung

n = jumlah sampel

Atribut yang numerik dalam penelitian ini, yaitu jumlah pemilih. Mean merupakan nilai rata-rata yang didapatkan dari hasil penjumlahan seluruh nilai dari masing-masing data, lalu dibagi dengan banyaknya data yang ada. Sedangkan standar deviasi merupakan nilai statistik yang dimanfaatkan untuk menentukan bagaimana sebaran data dalam sampel, serta seberapa dekat titik data individu ke mean atau rata-rata dari nilai sampel.

d. Menghitung Probabilitas, Selanjutnya, menghitung nilai probabilitas dari tiap atribut yang digunakan. Adapun rumus yang digunakan untuk menghitung probabilitas adalah :

$$P(C_i) = \frac{N_c}{N}$$

Keterangan :

$P(C_i)$ = Probabilitas kelas

N_c = Jumlah data dari kelas

N = Data pada kelas

Probabilitas merupakan suatu nilai yang digunakan untuk mengukur tingkat terjadinya suatu kejadian yang acak. Kata probabilitas ini sendiri sering disebut dengan peluang atau kemungkinan suatu atribut akan terjadi.

e. Menghitung nilai gaussian, yaitu menghitung nilai probabilitas untuk fitur data testing yang mempunyai data numerik/angka. Adapun persamaan dalam mencari nilai distribusi gaussian adalah:

$$P = (X_i = x_i | Y = y_j) = \frac{1}{\sqrt{2\pi\sigma_{ij}^2}} x e^{-\frac{(x_i - \mu_{ij})^2}{2\sigma_{ij}^2}}$$

Keterangan :

P = Peluang

X_i = Nilai atribut ke-i

Y_j = kelompok ke-j

σ_{ij}^2 = variansi kelompok ke-j

Pengujian data testing, Dari nilai probabilitas tiap atribut, data testing akan diuji berdasarkan model prediksi yang telah kita buat. Untuk data testing, disini menggunakan data yang didapatkan dari hasil kuesioner warga daerah kab.bone tentang partai yang akan dipilih. Untuk pengujian data testing ini akan menghasilkan prediksi partai pemenang dari tiap kecamatan.

4. HASIL DAN PEMBAHASAN

A. Data Awal

Data Awal didapatkan dari hasil pemilih umum DPRD Kab/Kota Pada periode tahun 2019, 2014, dan 2009. Data ini terdiri dari 1704 data, dimana terdiri dari partai, jumlah pemilih,kecamatan, tahun dan hasil menang atau kalah pada partai tersebut. Berikut merupakan tabel data hasil pemilu DPRD khususnya untuk tiap partai pemenang dari tiap kecamatan dari tiap periode tahun 2019, 2014, dan 2009.

Tabel 1. Data Awal

Partai	Jumlah Pemilih	Kecamatan	Tahun	Hasil
Partai Kebangkitan Bangsa	98	Palakka	2019	Kalah
Partai Gerakan Indonesia Karya	153	Tanete Riattang Barat	2019	Kalah
Partai Demokrasi Indonesia Perjuangan	129	Tanete Riattang Barat	2019	Kalah
.....
Partai Kebangkitan Bangsa	105	Tanete Riattang Barat	2014	Kalah
Partai Keadilan Sejahtera	101	Tanete Riattang Barat	2014	Kalah
Partai Golongan Karya	657	Palakka	2014	Menang
Partai Demokrat	143	Tanete Riattang Timur	2014	Kalah
.....
Partai Keadilan Sejahtera	28	Awangpone	2009	Kalah
Partai Amanat Nasional	126	Tellu Siatinge	2009	Kalah

Partai Perjuangan Indonesia Baru	18	Dua Boccoe	2009	Kalah
Partai Kedaulatan	23	Cenrana	2009	Kalah

B. Penerapan Metode Naive Bayes

Adapun beberapa tahapan dalam metode naive bayes, adalah :

1. Membaca data training, Disini, data training yang akan digunakan adalah data awal pada tabel 1. Yang akan digunakan seluruhnya untuk data training.
2. Menghitung jumlah dan probabilitas, Berdasarkan data yang digunakan, terdapat data numerik dan data kategorical. Adapun hasil dari perhitungan probabilitas untuk data kategorical adalah :

Tabel 2. Probabilitas Hasil

Probabilitas Hasil	Probabilitas
Kalah	0.956245
Menang	0.047535

Tabel 3. Probabilitas Partai

Probabilitas Partai	Kalah	Menang
Partai Kebangkitan Bangsa	0.046827	0.061728
Partai Gerakan Indonesia Karya	0.002465	0
Partai Demokrasi Indonesia Perjuangan	0.49908	0
Partai Golongan Karya	0.009242	0.814815
Partai Nasdem	0.33272	0
Partai Gerakan Perubahan Indonesia	0.16636	0
Partai Berkarya	0.01602	0.012346
Partai Keadilan Sejahtera	0.049291	0.012346
Partai Persatuan Indonesia	0.016636	0
Partai Pengusaha Dan Pekerja Indonesia	0.016636	0
Partai Peduli Rakyat Nasional	0.016636	0
Partai Barisan Nasional	0.016636	0
Partai Perjuangan Indonesia Baru	0.016636	0
Partai Kedaulatan	0.016636	0
Partai Persatuan Daerah	0.016636	0
Partai Pemuda Indonesia	0.016636	0
Partai Demokrasi Pembaruan	0.016636	0
Partai Karya Perjuangan	0.016636	0
Partai Matahari Bangsa	0.016636	0
Partai Penegak Demokrasi Indonesia	0.016636	0
Partai Demokrasi Kebangsaan	0.016636	0
Partai Republika Nusantara	0.016636	0
Partai Pelopor	0.016636	0
Partai Nasional Benteng Kerakyatan Indonesia	0.016636	0
Partai Bintang Reformasi	0.016636	0
Partai Patriot	0.016636	0
Partai Indonesia Sejahtera	0.016636	0
Partai Kebangkitan Nasional Ulama	0.016636	0
Partai Merdeka	0.016636	0
Partai Persatuan Nahdlatul Ummah Indonesia	0.016636	0
Partai Sarikat Indonesia	0.016636	0
Partai Nasional Indonesia Marhaenisme	0.007394	0
Partai Damai Sejahtera	0.003697	0
Partai Kasih Demokrasi Indonesia	0.003697	0
Partai Buruh	0.003697	0

Tabel 4. Probabilitas Kecamatan

Probabilitas Kecamatan	Kalah	Menang
Palakka	0.036352	0.037037
Tanete Riattang	0.036352	0.037037
Tanete Riattang Barat	0.036352	0.037037
Tanete Riattang Timur	0.036352	0.037037
Tonra	0.036969	0.037037
Mare	0.036969	0.037037
Sibulue	0.036969	0.037037
Barebbo	0.036969	0.037037
Cina	0.036969	0.037037
Ponre	0.036969	0.037037
Bontocani	0.038817	0.037037
Kahu	0.038817	0.037037
Kajuara	0.038817	0.037037
Salomekko	0.038817	0.037037
Libureng	0.038817	0.037037

Patimpeng	0.038817	0.037037
Awangpone	0.036352	0.037037
Tellu Siattinge	0.036352	0.037037
Cenrana	0.036352	0.037037
Lappariaja	0.036352	0.037037
Lamuru	0.036352	0.037037
Ulaweng	0.036352	0.037037
Amali	0.036352	0.037037
Tellulimpo	0.036352	0.037037
Bengo	0.036352	0.037037

Tabel 5. Probabilitas Tahun

Probabilitas Tahun	Kalah	Menang
2019	0.249538	0.333333
2014	0.182994	0.333333
2009	0.567468	0.333333

3. Menghitung nilai distribusi gaussian, Dalam menghitung distribusi gaussian yaitu hanya menghitung nilai probabilitas untuk fitur data testing yang type datanya adalah numerik pada data testing. Adapun hasilnya adalah :

Tabel 6. Distribusi Gaussian

Partai	Jumlah Pemilih	Kecamatan	Tahun	Hasil	Kalah	Menang
Partai Kebangkitan Bangsa	98	Palakka	2019	Kalah	0.040792	0.008482
Partai Gerakan Indonesia Karya	153	Tanete Riattang	2019	Kalah	0.04095	0.011687
Partai Demokrasi Indonesia Perjuangan	129	Tanete Riattang Barat	2019	Kalah	0.042986	0.010229
Partai Golongan Karya	171	Palakka	2019	Menang	0.037523	0.012828
Partai Nasdem	31	Tonra	2019	Kalah	0.023385	0.005336

4. Menghitung probabilitas akhir setiap class, Adapun probabilitas akhir dari setiap class dapat dilihat pada tabel berikut :

Tabel 7. Hasil Akhir

Partai	Jumlah Pemilih	Kecamatan	Tahun	Hasil	Kalah	Menang	Kalah	Menang	Prediksi
Partai Kebangkitan Bangsa	98	Palakka	2019	Kalah	0.040792	0.008482	6.61379e-05	9.21794e-07	Kalah
Partai Gerakan Indonesia Karya	153	Tanete Riattang	2019	Kalah	0.04095	0.011687	3.49445e-06	0	Kalah
Partai Demokrasi Indonesia Perjuangan	129	Tanete Riattang Barat	2019	Kalah	0.042986	0.010229	7.42809e-05	0	Kalah
Partai Golongan Karya	171	Palakka	2019	Menang	0.037523	0.012828	1.20073e-05	1.84027e-05	Menang
Partai Nasdem	31	Tonra	2019	Kalah	0.023385	0.005336	2.73964e-05	0	Kalah

5. Menghitung akurasi, berikut merupakan hasil dari confusion table :

Tabel 8. Akurasi

Prediksi	Menang	Kalah
Menang	2	0
Kalah	2	25
Akurasi	93%	

Dari tabel diatas, dapat dilihat bahwa akurasi didapatkan sebesar 93%.

5. KESIMPULAN DAN SARAN

Dari hasil dan pembahasan didapatkan kesimpulan bahwa untuk melakukan validasi data pemilih menggunakan NIK berdasarkan wawancara dan diimplementasikan di aplikasi untuk mendukung proses pengecekan validasi data pemilih sehingga dapat mengatasi permasalahan double data dan kesalahan input data.

Dan dari hasil dan pembahasan pula dapat disimpulkan bahwa hasil prediksi akan lebih akurat bergantung dari validasi data pemilih, didukung dengan data pemilu yang valid serta penerapan metode naive bayes sehingga didapatkan prediksi yang akurat. Hasil yang didapatkan dari metode naive bayes dalam memprediksi partai pemenang di KPU Kab.Bone adalah memperoleh akurasi dalam perhitungan manual sebesar 93% yang menandakan bahwa prediksi yang didapatkan sangat baik.

6. DAFTAR PUSTAKA

- [1] Fitriani, A. (2019). Penerapan Data Mining Menggunakan Metode Klasifikasi Naive bayes untuk Memprediksi Partisipasi Pemilihan Gubernur. JTAM (Jurnal Teori dan Aplikasi Matematika), 98-104.
- [2] Fauzi, A. (2018). Pengembangan Integrasi Sidalih Antara Pilwali Surabaya dan Pilgub Jawa Timur: Optimalisasi Pelayanan Publik KPU Kota Surabaya. JPSI (Journal of Public Sector Innovations), 3.
- [3] Tanthowi, P. U., Perdana, A., & Sukmajati, M. (2019). Tata Kelola Pemilu di Indonesia. Jakarta Pusat: Komisi Pemilihan Umum Republik Indonesia.
- [4] Santosa, S., & Yuliantara, R. (2017). Model Prediksi Pola Loyalitas Pelanggan Telekomunikasi Menggunakan Naive bayes Dengan Optimasi Particle Swarm Optimization. Jurnal Cyberku, 13(2), 7-7.
- [5] Widaningsih, S. (2019). Perbandingan Metode Data Mining Untuk Prediksi Nilai Dan Waktu Kelulusan Mahasiswa Prodi Teknik Informatika Dengan Algoritma C4, 5, Naive Bayes, Knn Dan Svm. Jurnal Tekno Insentif, 16-25.
- [6] Devita, R. N. (2018). Perbandingan kinerja metode naive bayes dan k-nearest neighbor untuk klasifikasi artikel berbahasa indonesia. JTIK (Jurnal Teknologi Informasi dan Ilmu Komputer), 427-434.
- [7] Putra, R. N., & Hasanuddin, H. (2017). Pemutakhiran Data Pemilih Pada Pemilihan Bupati dan Wakil Bupati Kabupaten Bengkalis di Kecamatan Mandau dan Kecamatan Bantan Tahun 2015. Doctoral Dissertation, Riau University.
- [8] Putra, R. A. (2019). Penerapan Naive Bayes Classifier dengan Gaussian Function Untuk Menentukan Kelompok UKT. Jurnal Informatika Global, 9(2).
- [9] Sabrina, D. F., & Saad, M. (2021). Keadilan Dalam Pemilu Berdasarkan Sistem Presidential Threshold. Widya Pranata Hukum: Jurnal Kajian dan Penelitian Hukum, 3(1), 15-37.
- [10] Sulastris, S., & Nugroho, Y. S. (2017). Penerapan Data Mining Untuk Prediksi Rating Penjualan Buku Menggunakan Metode Naive bayes. DutaCom, 12(2), 57-72.
- [11] Pasaribu, P. (2017). Peranan Partai Politik dalam Melaksanakan Pendidikan Politik. JPPUMA Jurnal Ilmu Pemerintahan dan Sosial Politik Universitas Medan Area, 51-59.
- [12] Sibuea, M. L., & Safta, A. (2017). Pemetaan Siswa Berpestrasi Menggunakan Metode K-Means Clustering. JURTEKSI (Jurnal Teknologi dan Sistem Informasi), 4(1), 85-92.
- [13] Supriadi, Y. (2017). Komunikasi Politik DPRD Dalam Meningkatkan Peran Legislasi Di Kota Bandung. Mediator: Jurnal Komunikasi, 25-36.
- [14] Hanifah, I., & Prastowo, B.N. (2016). Uji GPS Tracking Dalam Skala Transportasi Antar Kota. Ijeis, 6.2:175-186
- [15] Setyawati, E., Wijoyo, H., & Soeharmoko, N. (2020). Relational Database Management System (RDBMS). Kab.Banyumas, Jawa Tengah: pena persada.