**Prediksi Saham**

Muhammad Rizky Fauzi1\*, Zacky Anwar Makarim2

2Departemen Sistem Informasi, Universitas Komputer Indonesia, Indonesia

Email: \*[rizky.10519066@mahasiswa.unikom.ac.id](mailto:Ilman.10519069@mahasiswa.unikom.ac.id), [Zacky.10519079@mahasiswa.unikom.ac.id](mailto:Zacky.10519079@mahasiswa.unikom.ac.id)

Abstrak***—Prediksi saham adalah proses menentukan harga masa depan saham tertentu. Prediksi ini biasanya dilakukan oleh investor dan analis saham untuk membantu membuat keputusan investasi yang informatif. Prediksi saham bergantung pada berbagai faktor seperti kondisi ekonomi, industri, perusahaan, dan kebijakan pemerintah. Prediksi saham bisa dilakukan dengan berbagai metode, seperti analisis fundamental, analisis teknikal, dan pemodelan statistik. Analisis fundamental menekankan pada faktor-faktor eksternal seperti kondisi ekonomi dan perusahaan, sementara analisis teknikal memfokuskan pada data historis harga dan volume perdagangan saham. Pemodelan statistik menggunakan algoritme dan teknik pemodelan matematis untuk memprediksi harga saham. Prediksi saham merupakan hal yang penting bagi investor dan analis saham karena membantu mereka memahami bagaimana harga saham akan berkembang dan membuat keputusan investasi yang lebih informatif. Namun, perlu diingat bahwa prediksi saham tidak selalu akurat dan bisa terjadi perubahan dalam faktor-faktor yang mempengaruhi harga saham. Oleh karena itu, prediksi saham harus digunakan sebagai panduan saja dan bukan sebagai dasar utama untuk membuat keputusan investasi.***

*Kata kunci—Python, Prediksi, Saham, Matplotlib*

***Abstract—*** ***Stock prediction is the process of determining the future price of a particular stock. These predictions are usually made by stock investors and analysts to help make informed investment decisions. Stock predictions depend on various factors such as economic conditions, industry, companies and government policies. Stock predictions can be done using various methods, such as fundamental analysis, technical analysis, and statistical modeling. Fundamental analysis focuses on external factors such as economic and company conditions, while technical analysis focuses on historical data on stock prices and trading volume. Statistical modeling uses algorithms and mathematical modeling techniques to predict stock prices. Stock predictions are important to investors and stock analysts because they help them understand how stock prices will develop and make more informed investment decisions. However, keep in mind that stock predictions are not always accurate and factors that affect stock prices may change. Therefore, stock predictions should be used as a guide only and not as the primary basis for making investment decisions..***

*Keywords—Python, Predictions, Stocks, Matplotlib*

# Pendahuluan

Cristiano Ronaldo lahir di Funchal, Madeira, Portugal pada 5 Februari 1985, seorang bungsu dari empat bersaudara. Ibunya adalah Maria Dolores dos Santos Aveiro merupakan seorang juru masak dan ayahnya adalah Jose Dinis Aveiro merupakan seorang tukang kebun. Mereka harus bekerja keras untuk memenuhi kebutuhan keluarga [1]. Cristiano kemudian menikmati kegemilangan selama satu musim bersama Sporting sebelum hijrah ke klub terbesar di dunia, Manchester United. Di usia 18, dia belum melakukan apa-apa namun mendapat jersey ikonik nomor 7, yang sebelumnya digunakan oleh pemain-pemain hebat seperti George Best, Eric Cantona dan David Beckham."Hanya sedikit pemain yang dianggap sebagai George Best baru tetapi untuk pertama kali, ini terasa bagai sebuah pujian," tutur Best, figur ikonik si Setan Merah pada 2004. Langsung berhadapan dengan ekspektasi setinggi langit, orang biasa dipastikan bakal kesulitan tetapi bagi Cristiano muda hal tersebut adalah motivasi ekstra. Tentu saja kemasyhuran Cristiano ini ditandai oleh pencapaian-pencapaian hebat di atas lapangan, dia berhasil mempertahankan kemampuan top dan selalu gemilang di momen-momen terbesar, seperti Clasico melawan Barcelona atau mengambil tendangan penalti penting di final Liga Champions kontra Atletico Madrid. Ketika Real membutuhkan Ronaldo, Ronaldo memberi jawaban. Ronaldo bahkan mereformasi diri sendiri, berganti posisi lebih ke dalam menjadi penyerang tengah dengan kualitas luar biasa dan menjadi pencetak gol terbanyak sepanjang sejarah klub melewati torehan legenda seperti Ferenc Puskas, Alfredo Di Stefano dan Raul. Di tengah-tengah atmosfer hebat di Madrid yang bisa membuat pemain-pemain legendaris hancur karena tekanan, Cristiano justru menjulang dengan menciptakan standar emas yang harus diburu oleh pemain lainnya. Sebagai tambahan dari empat Ballon d'Or lainnya, Ronaldo menyambar sederet trofi bergengsi termasuk empat Liga Champions, tiga di antaranya didapatkan secara beruntun, sebelum mengambil pilihan bergabung ke Juventus demi sebuah tantangan baru pada musim panas 2018. Dia merupakan Mega bintang sepakbola dunia dari negara Portugal yang sekarang bermain untuk salah klub raksasa liga Inggris yaitu, Manchester United sebagai penyerang dan memperkuat timnas Portugal yang tak lain tak bukan ialah negara kelahirannya sendiri. Cristiano Ronaldo sendiri merupakan idola dari berbagai pemain muda yang kini bermunculan dan menjelma sebagai calon Mega bintang sepakbola dunia yang baru setelah era Cristiano Ronaldo dan Lionel Messi serta pemain bintang dunia lainnya yang karirnya usai.



*Gambar 1. Cristiano Ronaldo*

Pengaruh Cristiano Ronaldo tidak sampai disitu, bukti nyata bahwa pengaruh Cristiano Ronaldo memang memberikan pengaruh yang cukup besar ialah Ketika kepindahannya menuju salah satu klub sepakbola besar di Italia, yaitu Juventus, tak main-main, jumlah pengikut media sosial Juventus meningkat berkat kepindahan Cristiano Ronaldo dari Real Madrid ke Juventus. [3]

Tujuan dari penelitian ini adalah diantaranya ialah membandingkan kemampuan Cristiano Ronaldo dari setiap klub sepakbola yang dibela agar dapat menjadi acuan bagi para generasi muda pesepakbola. Metode penelitian yang digunakan dalam penelitian ini adalah metode *Naive Bayes* yang merupakan salah satu metode dalam pembelajaran mesin yang dikemukakan oleh seorang ilmuwan asal Ingggris yaitu, Thomas Bayes.

# PENELITIAN TERDAHULU

## Penelitian Terdahulu

Penelitian terdahulu menampilkan beberapa hasil penelitian yang telah dilakukan oleh para peneliti sebelumnya yang sekiranya masih berkait dengan tema analisa yang dilakukan oleh penulis dari judul analisa diatas. Dibawah ini merupakan hasil penelitian terdahulu yang dilakukan oleh para peneliti.

Yang pertama dari penelitian Haditsah Annur dari Universitas Ichsan Gorontalo yang berjudul “Klasifikasi Masyarakat Miskin Menggunakan Metode Naive Bayes” mengklasifikasikan berdasarkan data penduduk miskin yang diperoleh dari Kecamatan Tibawa dengan menggunakan teknik data mining dengan metode *Naive Bayes* dan hasilnya diperoleh tingkat akurasi 73% atau termasuk dalam kategori Baik. Sedangkan nilai Precision sebesar 92% dan Recall sebesar 86%. [4]

Yang kedua terdapat penelitian Solihati, T. I., Nurhidayanti, N., & Kania, R. (2022) dengan judul penelitian “Implementasi Data Mining untuk Evaluasi Kinerja Penelitian Mahasiswa Universitas Banten Jaya Dengan Menggunakan Algoritma Naïve Vayes Classifier” dimana fokus penelitian ini untuk mengevaluasi kinerja penelitian mahasiswa dari tahun 2019-2020 di Unversitas Banten Jaya dengan menggunakan data mining. Data mining memiliki tujuan untuk memanfaatkan data dalam basis data dengan mengolahnya sehingga menghasilkan informasi baru yang berguna. Pada penelitian ini menggunakan algoritma Naïve Bayes Classifier (NBC) untuk membentuk tabel probabilitas sebagai dasar proses klasifikasi kinerja penelitian mahasiswa. Atribut yang digunakan terdapat 5 macam dengan masing-masing memiliki value ya dan tidak. Berdasarkan hasil klasifikasi NBC pada tools weka diperoleh kesimpulan bahwa metode NBC berhasil mengklasifikasikan 50 data dari 64 data yang diuji. Sehingga metode NBC berhasil memprediksi artikel mahasiswa yang sudah terpublish dengan presentase akurasi sebesar 78,125% sementara presentase untuk incorrectly classified instances sebesar 21,875% . Hasil dari NBC dapat digunakan untuk mengklasifikasikan kinerja penelitian mahasiswa tahun selanjutnya dan data penelitian ini dapat dijadikan sebagai data training. [5]

Yang ketiga ada penelitian dari Darmawan, Muhammad Farhan Khoirulamri (2022) dengan judul penelitian “IMPLEMENTASI ALGORITMA NAIVE BAYES CLASSIFIER UNTUK KLASIFIKASI BIDANG DOSEN BERDASARKAN JUDUL PENELITIAN MAHASISWA”. dari Universitas AMIKOM Yogyakarta dimana Masalahnya, masih ada mahasiswa yang sering mengeluhkan dosen pembimbing tidak sesuai dengan bidangnya sehingga mengakibatkan sulitnya dalam proses bimbingan. Dalam penelitian kali ini, data yang digunakan diperoleh dari Universitas AMIKOM Yogyakarta Program Studi Informatika tahun 2014 hingga 2016 yang terdiri dari 1598 data. Kemudian, dataset tersebut akan diolah melalui proses Preprocessing, pembobotan kata dan diklasifikasikan menggunakan algoritma naïve bayes classifier apakah dataset tersebut sudah sesuai dengan bidang dosen terdiri dari Data Science, Multimedia, Networking, Software Engineer, dan Game. Setelah itu dataset dibagi menjadi data training dan data test sebanyak 9:1 (90% data training dan 10% data test), Berdasarkan hasil penelitian yang telah dilakukan, hasil akurasi yang diperoleh dari klasifikasi menggunakan algoritma naïve bayes classifier sebesar 88.29%. [6]

Lalu yang keempat ada penelitian dari Nani Purwati dari Universitas Bina Sarana Informatika yang berjudul “Analisa Pengaruh Iklan tanpa Label Harga pada media sosial Menggunakan Algoritma Naive Bayes”. Penelitian ini bertujuan menganalisa keputusan calon konsumen terhadap iklan tanpa label harga. Metode penelitian yang digunakan pada penelitian ini adalah algoritma klasifikasi data mining Algoritma Naive Bayes. Sedangkan, untuk metode penerapan data mining menggunakan metodologi Knowledge Discovery of Database (KDD). Adapun, sumber data dalam penelitian ini diperoleh dari hasil sebaran quisioner terhadap 203 responden pengguna media sosil aktif dengan atribut penelitian berdasarkan gender, attention, interest , search dan Action. Setelah melakukan tahapan proses KDD pada data mining menggunakan algoritma Naive Bayes diperoleh akurasi sebesar 97,6%. Berdasarkan hasil penelitian tersebut, dapat disimpulkan bahwa iklan tanpa label harga pada media sosial dapat mempengaruhi keputusan pembelian. [7]

Kemudian, yang kelima terdapat penelitian Damayanti, Widya (2022) dengan judul “Klasifikasi Pendapat Masyarakat Terhadap Penggunaan Vaksin Dalam Mengantisipasi Covid-19 Menggunakan Teknik Sentimen Analisis Berbasis Naive Bayes” dari Universitas Muhammadiyah Jember. Penelitian ini berisi tentang analisis opini masyarakat terhadap vaksinasi menggunakan metode berbasis Naive Bayes yang dihimpun dari timeline Twitter sejak tanggal 23 Februari sampai 9 Maret 2021. Jumlah opini yang berhasil dihimpun dan dijadikan data dalam penelitian ini adalah 1000 data tweet yang dibagi menjadi 200 tweet data uji validasi dan 800 tweet data latih. Penelitian ini juga menggunakan Chi-Square sebagai metode seleksi fitur. Algoritma yang digunakan dalam penelitian ini adalah Naive Bayes dan Multinomial Naive Bayes . Dari hasil pengukuran pada pengujian yang dilakukan, metode Naive Bayes memperoleh nilai akurasi sebesar 66%, presisi sebesar 79% dan recall sebesar 62%. Sedangkan metode Multinomial Naive Bayes memperoleh nilai akurasi sebesar 71%, presisi sebesar 81% dan recall sebesar 68%. Berdasarkan hasil yang diperoleh, algoritma Multinomial Naive Bayes memiliki performa yang lebih baik daripada Naive Bayes. [8]

Setelah itu, terdapat penelitian yang dilakukan oleh Dianati Duei Putri, Gigih Forda Nama, dan Wahyu Eko Sulistiono dari Universitas Lampung dengan judul penelitian “Analisis Sentimen Kinerja Dewan Perwakilan Rakyat (DPR) Pada Twitter Menggunakan Metode Naive Bayes Classifier”. Dalam penelitian ini akan dilakukan analisis sentimen masyarakat terhadap kinerja Dewan Perwakilan Rakyat (DPR) yang diungkapkan melalui media sosial twitter. Ada beberapa tahap untuk melakukan analisis sentimen , yaitu pengumpulan data (crawling), preporcessing data yang terdiri dari proses cleaning data, tokenization, stop remova dan case folding, splitting data dan klasifikasi data menggunakan metode Naive Bayes Classifier. Penelitian ini menggunakan sebanyak 1546 data tweet. Hasil dari penelitian ini didapatkan bahwa DPR mendapatkan 95 tweet positif dengan polaritas 0.75 atau 75% sentimen positif, 693 tweet netral dengan polaritas 0.79 atau 79% sentimen netral dan 758 tweet negatif dengan polaritas 0.82 atau 82% sentimen negatif dengan accuracy score 0.8 atau 80% berdasarkan data testing sebanyak 20%. [9]

Kemudian ada penelitian yang dilakukan oleh Risna Alfiani dan Yuyun Umaidah dari Teknik Informatika, Universitas Singaperbangsa Karawang dengan judul penelitian “Prediksi Kepuasan Customer Terhadap Performance Terapis Baby Massage Menggunakan Algoritma Naive Bayes”. Penelitian ini dilakukan dengan tujuan untuk mengetahui besarnya tingkat akurasi klasifikasi yang dihasilkan dari prediksi kepuasan pelanggan terhadap performa terapis sebuah baby massage dengan menggunakan algoritma Naïve Bayes Penelitian ini menggunakan data penilaian kepuasan pelanggan yang diambil dengan metode penyebaran kuesioner kepada pelanggan sebuah baby massage di kota Surakarta. Data tersebut akan diolah dan dianalisis menggunakan aplikasi RapidMiner untuk mendapatkan hasil yang dibutuhkan. Hasil dari penelitian ini adalah prediksi pola kepuasan pelanggan baby massage yang memetakan penilaian pelanggan menjadi dua label yaitu puas dan tidak puas. Hasil yang diperoleh dari pengolahan data menggunakan algoritma Naive Bayes menunjukkan hasil prediksi dengan nilai akurasi sebesar 98,21% yang berarti algoritma Naive Bayes cocok untuk memprediksi kepuasan pelanggan terhadap performa terapis baby massage. [10]

## Landasan Teori

1. Definisi Data Mining

Data mining adalah metode dalam ilmu komputer yang biasa digunakan dalam proses pencarian knowledge. Tahapan di dalamnya berguna untuk mencari pola-pola tertentu dari data yang ada pada database. Biasanya, metode ini banyak ditemukan pada bidang machine learning dan statistika.

Pada awalnya, metode penambangan data dikembangkan karena kompleksitas kerja komputer yang semakin meningkat. Namun, disinilah keuntungan adanya data mining adalah proses pengumpulan dan seleksi data yang lebih praktis.

Sebagai definisi lain, ada beberapa pengertian data mining menurut para ahli yang juga patut dipelajari. Simak penjelasan secara singkat di bawah ini.

Larose (2006), mengartikan data mining adalah sebuah proses menemukan sesuatu bermakna dengan memilah data melalui repository dengan bantuan teknologi sosialisasi pola, statistik, serta matematika.

Berry, data mining adalah sebuah aktivitas analisa data yang memiliki jumlah besar demi menemukan pola (pattern) dan aturan (rule) yang berguna.

Pramudiono (2006), menyampaikan bahwa data mining adalah proses analisa yang dilakukan secara otomatis pada data yang kompleks dan berjumlah besar untuk memperoleh sebuah pola atau kecenderungan yang umumnya tidak disadari. [11]

1. Metode Data Mining

Metode data mining adalah cara yang diterapkan namun perlu disesuaikan dengan tujuan penggunanya. Ada beberapa pembagian metode data mining berikut yang bisa ketahui.

Classification

Klasifikasi data mining adalah sebuah proses menemukan definisi kesamaan karakteristik dalam suatu kelompok atau kelas (class). Klasifikasi data mining menjadi salah satu metode yang paling umum untuk digunakan. Metode ini dilakukan bertujuan untuk memperkirakan kelas dari suatu objek yang labelnya belum diketahui.

Association

Metode data mining yang kedua adalah Market basket analysis (analisa keranjang pasar) atau association. Berhubungan dengan pemasaran, metode ini bertujuan untuk mengidentifikasi produk yang sering dibeli bersamaan oleh pelanggan.

Misalnya beberapa pelanggan akan membeli snack dan minuman kemasan bersamaan. Dengan begitu perusahaan lebih mudah mengetahui kalau kedua barang tersebut sering dibeli bersamaan.

Clustering

Clustering data mining adalah nama lain untuk metode segmentation. Tujuan dari segmentasi pada metode data mining adalah mengelompokkan suatu class ke dalam beberapa segmen berdasarkan atribut yang ditentukan. Penentuan atribut harus sesuai kesamaan yang dimiliki beberapa class tadi.

Regression

Metode keempat dari data mining adalah regression. Metode yang satu ini sedikit mirip dengan klasifikasi data mining. Bedanya, regresi merupakan metode yang bertujuan untuk mencari pola nilai numerik, bukan kelas. Hasil dari metode regression adalah sebuah fungsi sebagai penentu hasil yang didasarkan nilai dari input.

Forecasting

Forecasting data mining adalah metode yang digunakan untuk memprediksi nilai yang akan dicapai pada satu periode. Dengan menggunakan teknik ini, noise data dan nilai pada periode sebelumnya dijadikan dasar bahan prediksi.

Sequencing

Sequence adalah sebuah urutan peristiwa. Tidak jauh dari namanya, metode sequence analysis berfungsi untuk mencari sebuah pola pada serangkaian kejadian atau sequence. Contoh sederhananya dalam kehidupan sehari-hari adalah ketika makan. Rangkaian yang umum dilakukan adalah mengambil piring, sendok dan garpu, kemudian baru mengambil lauk pauk. Pola seperti itu lah yang berusaha dicari tahu melalui sequencing.

Descriptive

Metode data mining yang satu ini bertujuan untuk memahami lebih dalam mengenai data-data yang masuk dalam pengamatan. Hasil akhirnya adalah mengetahui perilaku dari data itu sendiri. [11]

1. Tahapan Data Mining

Pembersihan data (data cleaning)

Dalam proses menambang emas, tahap awal yang dilakukan adalah dengan mencari batu atau lumpur yang sekiranya bisa diolah menjadi emas yang bagus. Dalam proses tersebut, pasti menemukan satu atau bahkan mencapai puluhan batu yang kurang layak untuk diproses. Sehingga batu-batu itu perlu dibuang.

Begitu juga penerapannya dalam tahapan data mining. Sebelum diproses dan dibentuk menjadi sebuah knowledge, data yang ada harus dibersihkan terlebih dahulu. Jika ada yang mengandung eror, maka data-data tersebut harus dibuang. Sehingga tersisa data yang ‘bagus’ untuk diolah dalam tahap selanjutnya.

Integrasi data (data integration)

Langkah kedua dalam tahapan data mining adalah integrasi data. Setelah menemukan batu-batu yang cocok, selanjutnya penambang akan mulai mengkombinasikan untuk dijadikan batangan emas atau bentuk emas lainnya. Dalam data mining, data yang berhasil dibersihkan juga akan diintegrasi.

Transformation

Seperti bentuk emas yang berbeda-beda, maka proses pembentukannya juga akan berbeda. Begitu juga dengan data, dalam tahap transformation data akan dipilih dan diubah formatnya agar sesuai dengan teknik atau metode yang dipakai. Pada tahap ini pula kualitas data mining akan terlihat.

Data mining

Tahapan data mining selanjutnya adalah proses penambangan data itu sendiri. Perlu penentuan metode penambangan yang tepat.

Evaluasi pola (pattern evaluation)

Setelah bahan mentah emas selesai diolah, maka perlu diuji juga apakah sudah sesuai dengan standar kualitasnya atau belum. Baru setelah itu bisa didistribusikan ke toko emas dan di-display kepada pelanggan.

Setelah selesai melakukan proses data mining, pola-pola yang dihasilkan dari proses tersebut perlu untuk dievaluasi. Tujuan dari dilakukannya evaluasi adalah untuk menguji hipotesis awal. Setelah teruji data bisa dipresentasikan kepada pengguna. [11]

1. Manfaat Data Mining

Secara teknis, metode ini memang tidak lepas dari ilmu komputer dan kecerdasan buatan. Tetapi ternyata manfaatnya bisa dirasakan dalam berbagai bidang lainnya, termasuk bisnis dan pemasaran. Berikut sejumlah manfaat data mining.

Mengetahui tren

Metode untuk memprediksi keputusan bisnis di masa depan

Mengetahui produk yang dibeli bersamaan

Mengamati perilaku konsumen

Model sebagai sarana menyusun strategi peningkatan penjualan. [11]

1. Penjelasan Naive Bayes

Naive Bayes adalah metode klasifikasi data berdasarkan probabilitas yang bisa saja terjadi di masa depan. Metode ini dianggap simpel dan efektif untuk diterapkan dalam analisis perusahaan.

Naive Bayes adalah metode yang cocok untuk klasifikasi biner dan multiclass. Metode yang juga dikenal sebagai Naive Bayes Classifier ini menerapkan teknik supervised klasifikasi objek di masa depan dengan menetapkan label kelas ke instance/catatan menggunakan probabilitas bersyarat. Probabilitas bersyarat adalah ukuran peluang suatu peristiwa yang terjadi berdasarkan peristiwa lain yang telah (dengan asumsi, praduga, pernyataan, atau terbukti) terjadi.

Istilah supervised merujuk pada klasifikasi training data yang sudah diberi label dengan kelas. Misalnya, sebuah transaksi penipuan telah ditandai sebagai data transaksional. Kemudian, jika Anda ingin mengklasifikasikan transaksi di masa depan menjadi fraudulent/non-fraudulent (penipuan/non-penipuan), maka jenis klasifikasi itu akan disebut sebagai supervised.

Model machine learning yang diterapkan pada program tersebut menggunakan teorema Bayes yang dirumuskan sebagai berikut:

P(A│B) = P(B│A)P(A)P(B)

Keterangan:

P(A│B) : Probabilitas A terjadi dengan bukti bahwa B telah terjadi (probabilitas superior)

P(B│A) : Probabilitas B terjadi dengan bukti bahwa A telah terjadi

P(A) : Peluang terjadinya A

P(B) : Peluang terjadinya B

Dalam istilah yang lebih sederhana, Teorem Bayes adalah rumus matematika sederhana untuk menemukan probabilitas ketika kita mengetahui probabilitas tertentu lainnya. [12]

1. Tujuan dan Contoh Penerapan Naive Bayes

Tujuan metode ini adalah mengklasifikasikan probabilitas berdasarkan pembelajaran mesin atas probabilitas lain. Teknik ini dapat dimanfaatkan untuk mengevaluasi banyak hal, seperti misalnya prakiraan cuaca. Prakirawan cuaca dapat memprediksi apakah besok akan cerah, hujan, atau berangin berdasarkan suhu, kelembapan, tekanan, dan lain-lain.

Media sosial juga dapat menampilkan konten-konten yang kemungkinan Anda sukai berdasarkan konten yang telah Anda lihat sebelumnya. Selain itu, metode ini juga efektif untuk menganalisis ancaman fraud/penipuan. Dengan memprediksi ancaman-ancaman yang muncul, Anda dapat mengambil solusi lebih awal untuk menghindari hal-hal yang tidak diinginkan.

Metode ini juga dapat diterapkan petugas bank untuk memprediksi apakah pemohon pinjaman akan menjadi defaulter/non-defaulter bank berdasarkan data jumlah pinjaman, angsuran bulanan, tenor, jumlah tunggakan, pendapatan tahunan, dan rasio utang. Berdasarkan pelabelan yang di-input, petugas bank mampu membuat keputusan apakah akan memberikan pinjaman pada pemohon dan berapa banyak kredit serta suku bunga yang kira-kira dapat diterima oleh pemohon.

Pada sektor medis dan kesehatan, dokter dapat memprediksi keberhasilan pengobatan atau kondisi pasien berdasarkan data seperti tekanan darah, kadar hemoglobin, kadar gula darah, nama obat yang diberikan, jenis pengobatan, dan sebagainya. Melalui data ini, dokter dapat menentukan metode pengobatan dan alternatif selanjutnya. [12]

1. Tipe Naive Bayes Classifier

Metode Naive Bayes digolongkan menjadi beberapa tipe berdasarkan fungsinya. Berikut ini penjelasannya.

MULTINOMIAL NAIVE BAYES

Salah satu tipe metode Naive Bayes adalah Multinomial yang sebagian besar digunakan untuk mengklasifikasi kategori dokumen. Sebuah dokumen dapat dikategorikan bertema olahraga, politik, teknologi, atau lain-lain berdasarkan frekuensi kata-kata yang muncul dalam dokumen.

2. BERNOULLI NAIVE BAYES

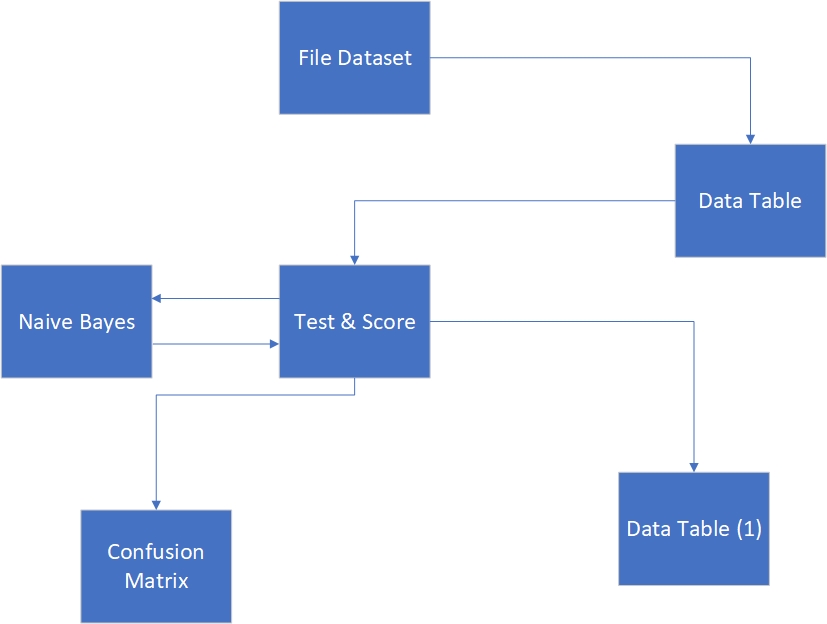
Tipe ini mirip dengan tipe Multinomial, namun klasifikasinya lebih berfokus pada hasil ya/tidak. Prediktor yang di-input adalah variabel boolean. Misalnya, prediksi atas sebuah kata muncul dalam teks atau tidak.

3. GAUSSIAN NAIVE BAYES

Distribusi Gaussian adalah asumsi pendistribusian nilai kontinu yang terkait dengan setiap fitur berisi nilai numerik. Ketika diplot, akan muncul kurva berbentuk lonceng yang simetris tentang rata-rata nilai fitur. [12]

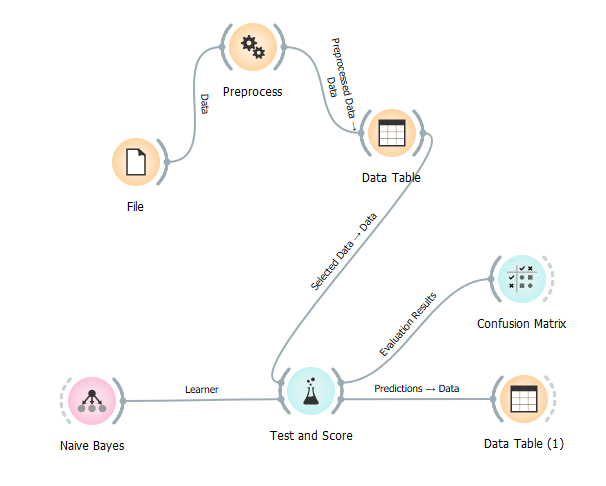
# MetodE

Sebelum mengimplementasikan metode Naive Bayes ke dalam aplikasi Orange Data Mining, maka pada gambar di bawah ini akan diilustrasikan bagaimana proses metode Naive Bayes.



*Gambar 2. Ilustrasi Metode Naive Bayes*

Dari ilustrasi pada gambar diatas, terdapat  *file* yang mencakup *dataset*, lalu *preprocessing*, tabel data, metode Naive Bayes, *test & score*,  *confusion matrix* yang menjadi komponen praktikum metode  *Naive Bayes* yang berguna dalam melakukan prediksi yang dilakukan.



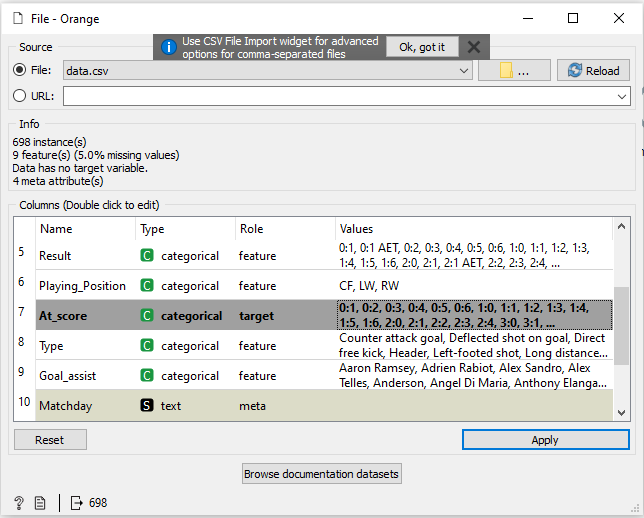
*Gambar 3. Implementasi Orange*

Gambar diatas meruapakan implementasi algoritma  *Naïve*  *Bayes* pada aplikasi orange dimana aplikasi tersebut berguna dalam melakukan pengolahan data mining.

# Hasil

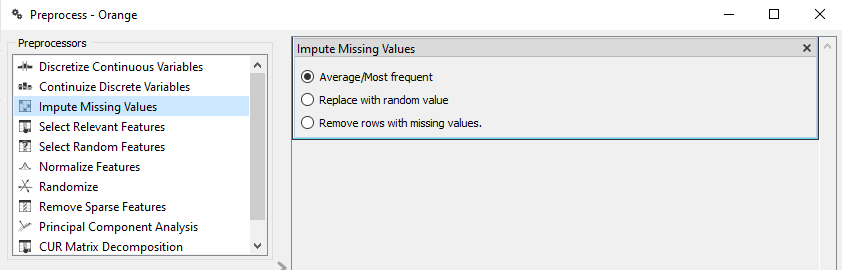
## Dataset

Dataset ini berisi semua statistik dari semua gol klub Cristiano Ronaldo dos Santos Aveiro.



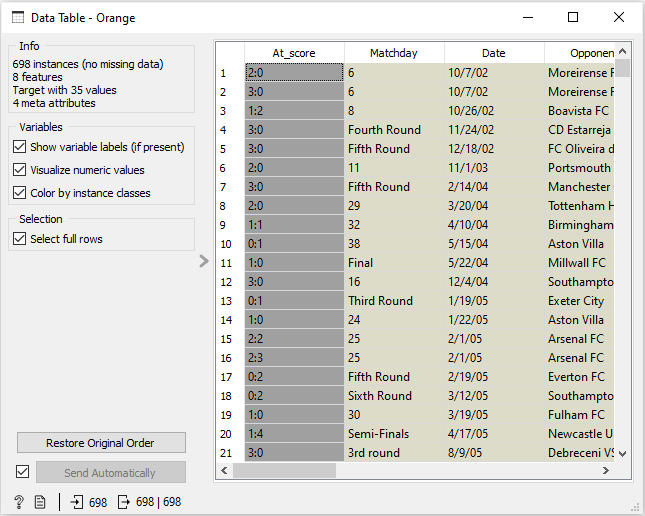
*Gambar 3. Dataset*

Didalam dataset tersebut terdapat  *score* yang dijadikan target data yang dimisalkan sebagai kunci atas pengolahan data yang dilakukan.



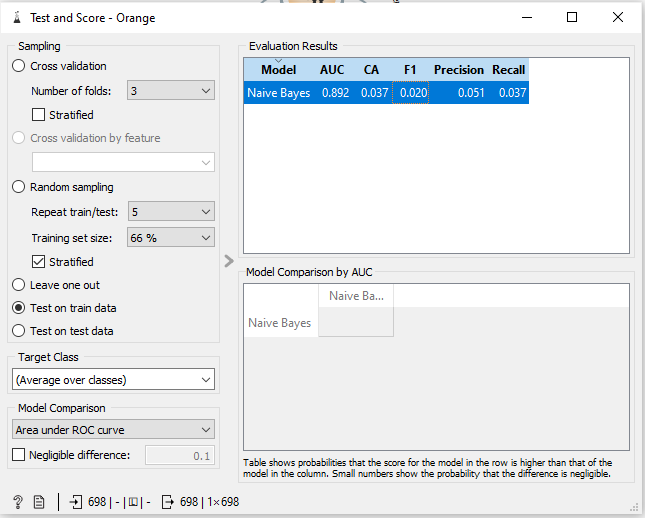
*Gambar 4. Preprocessing*

Preprocessing digunakan di dalam proses rata-rata



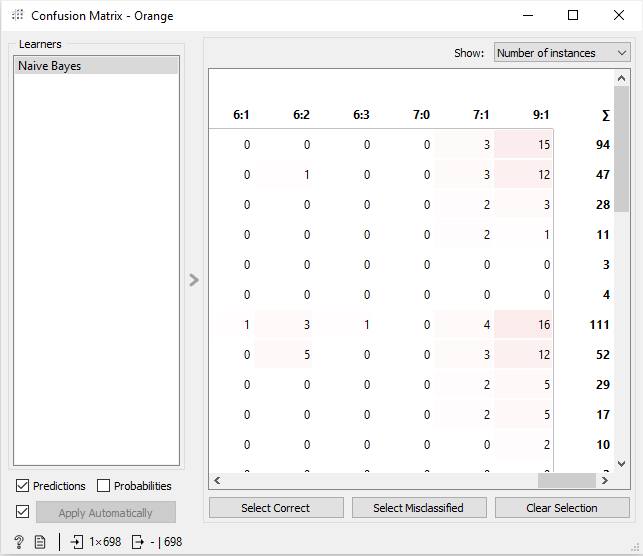
*Gambar 5. Data Table*

Tabel data yang berisi variabel



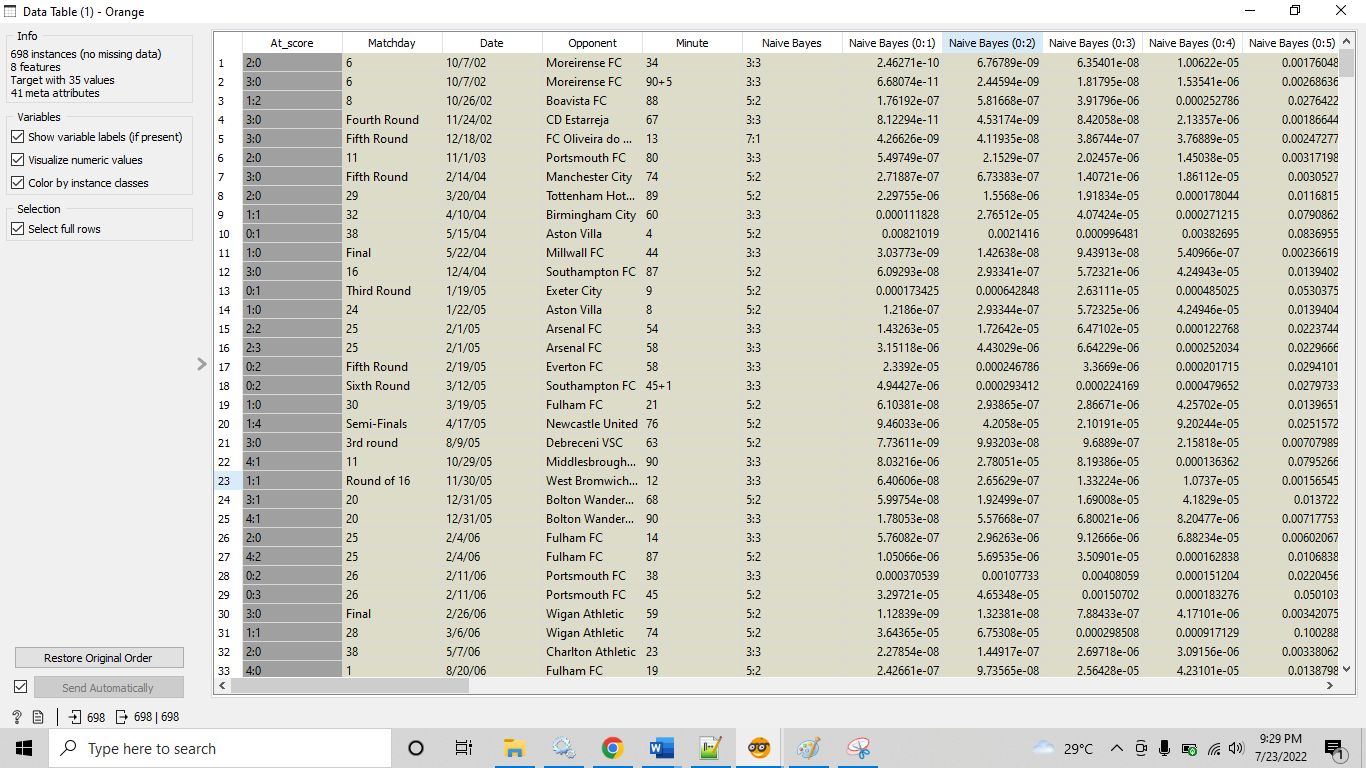
*Gambar 6. Test & Score*

Hasil pengolahan data dari metode  *Naïve Bayes* yang terdapat *Precision* dan *Recall*



*Gambar 7. Confusion Matrix*

Confusion Matrix dalam algoritma *Naïve Bayes*



*Gambar 8. Data Table*

Tabel data yang telah menggunakan metode *NaiveBayes*

# Kesimpulan

Kesimpulan dari prediksi saham adalah bahwa prediksi saham merupakan proses penting bagi investor dan analis saham dalam membuat keputusan investasi yang informatif. Prediksi saham didasarkan pada berbagai faktor seperti kondisi ekonomi, industri, perusahaan, dan kebijakan pemerintah, dan bisa dilakukan dengan berbagai metode seperti analisis fundamental, analisis teknikal, dan pemodelan statistik.

Walaupun prediksi saham bisa membantu memahami bagaimana harga saham akan berkembang, perlu diingat bahwa prediksi saham tidak selalu akurat dan bisa terjadi perubahan dalam faktor-faktor yang mempengaruhi harga saham. Oleh karena itu, prediksi saham harus digunakan sebagai panduan saja dan bukan sebagai dasar utama dalam membuat keputusan investasi.

Secara keseluruhan, prediksi saham merupakan hal yang penting dalam dunia investasi saham, namun perlu dilakukan dengan hati-hati dan menggabungkan informasi dari berbagai sumber untuk membuat keputusan investasi yang tepat.

##### VI. PEngakuan

Kami mengucapkan terima kasih kepada Universitas Komputer Indonesia yang telah membantu kami dalam penulisan makalah ini.

##### VII. Referensi

1. Goal, O. (n.d.). *Cristiano Ronaldo perjalanan panjang*. Soccer News, Live Scores, Results & Transfers. Retrieved July 18, 2022, from https://www.goal.com/story/cristiano-ronaldo-perjalanan-panjang/index.html
2. *Pengaruh Positif Cristiano Ronaldo, Pemain Manchester ... - tribunnews.com*. (n.d.). Retrieved July 17, 2022, from https://www.tribunnews.com/superskor/2021/09/14/pengaruh-positif-cristiano-ronaldo-pemain-manchester-united-tidak-berani-makan-sembarangan