**Analisis Data Bank Konvesional & Bank Syariah Menggunakan   
Naive Bayes pada Python**

Rikad Fauzi Alawi

Universitas Komputer Indonesia

Jln. Dipatiukur, Bandung

E-mail : rikad@email.unikom.ac.id

**A****BSTRAK**

Analisis Data adalah teknik untuk menemukan sesuatu yang menarik, berpotensi, berguna dan berarti dari sejumlah besar data. Analisis ini biasanya diaplikasikan diberbagai bidang seperti pemasaran, keuangan, perbankan, manufaktur dan telekomunikasi. Analisi Finansial bank adalah salah – satu contoh pentingnya dari sebuah analisis data sehingga kita bisa mempelajari tingkat kebangkrutan dari sebuah bank. kebangkrutan adalah untuk menentukan apakah suatu organisasi atau perusahaan keuangan akan bangkrut atau tidak. Kebangkrutan atau financial distress adalah suatu situasi yang timbul ketika sebuah bisnis atau keuangan perusahaan tidak mampu memenuhi kewajiban keuangan atau kewajiban utang. Prediksi kebangkrutan adalah masalah klasifikasi biner, dengan dua kelas: Bangkrut atau tidak bangkrut. Berbagai komponen dari kebangkrutan adalah milik likuidasi, dijamin utang, pembayaran, utang-batas, dll. Tantangan pertama dalam mempelajari kebangkrutan memulai dengan keputusan pemilihan model prediksi yang cocok, salah satu algoritma untuk prediksi yang dapat digunakan adalah algoritma naïve bayes dan untuk melakukan prediksi dapat menggunakan environtment pada python.

**Kata kunci :** naive bayes, bank syariah, finansial distress, bank konvensional

1. **PENDAHULUAN**

Analisis Data adalah teknik untuk menemukan sesuatu yang menarik, berpotensi, berguna dan berarti dari sejumlah besar data. Analisis ini biasanya diaplikasikan diberbagai bidang seperti pemasaran, keuangan, perbankan, manufaktur dan telekomunikasi. Analisi Finansial bank adalah salah – satu contoh pentingnya dari sebuah analisis data sehingga kita bisa mempelajari tingkat kebangkrutan dari sebuah bank. kebangkrutan adalah untuk menentukan apakah suatu organisasi atau perusahaan keuangan akan bangkrut atau tidak. Kebangkrutan atau financial distress adalah suatu situasi yang timbul ketika sebuah bisnis atau keuangan perusahaan tidak mampu memenuhi kewajiban keuangan atau kewajiban utang. Prediksi kebangkrutan adalah masalah klasifikasi biner, dengan dua kelas: Bangkrut atau tidak bangkrut. Berbagai komponen dari kebangkrutan adalah milik likuidasi, dijamin utang, pembayaran, utang-batas, dll.

Prediksi kebangkrutan adalah masalah klasifikasi biner, dengan dua kelas: Bangkrut atau tidak bangkrut. Berbagai komponen dari kebangkrutan adalah milik likuidasi, dijamin utang, pembayaran, utang-batas, dll. Tantangan pertama dalam mempelajari kebangkrutan memulai dengan keputusan pemilihan model prediksi. Peningkatan jumlah perusahaan yang mengalami kebangkrutan telah dimulai sejak tahun 1930 setelah ekonomis depresi. Prediksi kebangkrutan bank telah luas daerah penelitian sejak akhir Tahun 1960-an. Tujuan dari prediksi kebangkrutan model untuk memprediksi apakah seorang calon mahasiswa baru yang mencakup individu dan perusahaan akan bangkrut atau tidak. Jika prediksi model tidak bisa melakukan dengan baik dan tepat, hal itu akan menyebabkan inti dari keputusan yang salah dan kemudian hal ini sangat mungkin untuk menyebabkan krisis keuangan dan kesusahan. Dengan demikian, akurat prediksi kebangkrutan adalah masalah yang sangat penting di bidang keuangan. Kemajuan teknologi informasi memungkinkan kita untuk mendapatkan berbagai informasi tentang beberapa risiko status perusahaan dalam banyak hal, seperti lembaga profesional, media massa, dll. Penyebab kegagalan bisnis dan kebangkrutan dapat menunjuk ke: ekonomi, keuangan, kelalaian, penipuan, bencana dan lain-lain. Faktor-faktor ekonomi yang mengarah ke kebangkrutan termasuk kelemahan dari industri dan lokasi miskin, sementara faktor-faktor keuangan termasuk moderat utang dan kekurangan modal. Memprediksi kegagalan perusahaan adalah sangat penting ilmu manajemen masalah.

Tantangan pertama dalam mempelajari kebangkrutan memulai dengan keputusan pemilihan model prediksi yang cocok, salah satu algoritma untuk prediksi yang dapat digunakan adalah algoritma naïve bayes dan untuk melakukan prediksi dapat menggunakan environtment pada python.

Algoritma Naive Bayes merupakan sebuah metoda klasifikasi menggunakan metode probabilitas dan statistik yang dikemukakan oleh ilmuwan Inggris Thomas Bayes. Algoritma Naive Bayes memprediksi peluang di masa depan berdasarkan pengalaman di masa sebelumnya sehingga dikenal sebagai Teorema Bayes. Ciri utama dari Naive Bayes Classifier ini adalah asumsi yg sangat kuat (naïf) akan independensi dari masing-masing kondisi / kejadian. Keuntungan penggunan metode ini yaitu hanya membutuhkan jumlah data pelatihan (training data) yang kecil untuk menentukan estimasi parameter yang diperlukan dalam proses pengklasifikasian. Karena yang diasumsikan sebagai variabel independent, maka hanya varians dari suatu variabel dalam sebuah kelas yang dibutuhkan untuk menentukan klasifikasi, bukan keseluruhan dari matriks kovarians.

Kelebihan Naïve Bayes

* Bisa dipakai untuk data kuantitatif maupun kualitatif
* Tidak memerlukan jumlah data yang banyak
* Tidak perlu melakukan data training yang banyak
* Jika ada nilai yang hilang, maka bisa diabaikan dalam perhitungan.
* Perhitungannya cepat dan efisien
* Mudah dipahami
* Mudah dibuat
* Pengklasifikasian dokumen bisa dipersonalisasi, disesuaikan dengan kebutuhan setiap orang
* Jika digunakan dalaam bahasa pemrograman, code-nya sederhana
* Bisa digunakan untuk klasifikasi masalah biner ataupun multiclass

Kekurangan Naïve Bayes

* Apabila probabilitas kondisionalnya bernilai nol, maka probabilitas prediksi juga akan bernilai nol
* Asumsi bahwa masing-masing variabel independen membuat berkurangnya akurasi, karena biasanya ada korelasi antara variabel yang satu dengan variabel yang lain
* Keakuratannya tidak bisa diukur menggunakan satu probabilitas saja. Butuh bukti-bukti lain untuk membuktikannya.
* Untuk membuat keputusan, diperlukan pengetahuan awal atau pengetahuan mengenai masa sebelumnya. Keberhasilannya sangat bergantung pada pengetahuan awal tersebut Banyak celah yang bisa mengurangi efektivitasnya
* Dirancang untuk mendeteksi kata-kata saja, tidak bisa berupa gambar

Python merupakan bahasa pemrograman tingkat tinggi yang diracik oleh Guido van Rossum. Python banyak digunakan untuk membuat berbagai macam program, seperti: program CLI, Program GUI (desktop), Aplikasi Mobile, Web, IoT, Game, Program untuk Hacking, dsb. Python juga dikenal dengan bahasa pemrograman yang mudah dipelajari, karena struktur sintaknya rapi dan mudah dipahami. Python mendukung multi paradigma pemrograman, utamanya; namun tidak dibatasi; pada pemrograman berorientasi objek, pemrograman imperatif, dan pemrograman fungsional. Salah satu fitur yang tersedia pada python adalah sebagai bahasa pemrograman dinamis yang dilengkapi dengan manajemen memori otomatis. Seperti halnya pada bahasa pemrograman dinamis lainnya, python umumnya digunakan sebagai bahasa skrip meski pada praktiknya penggunaan bahasa ini lebih luas mencakup konteks pemanfaatan yang umumnya tidak dilakukan dengan menggunakan bahasa skrip. Python dapat digunakan untuk berbagai keperluan pengembangan perangkat lunak dan dapat berjalan di berbagai platform sistem operasi.

Scikit adalah Salah satu library yg memudahkan untuk membuat machine learning pada python. Menggunakan library ini memudahkan pemrograman karena lebih sedikit baris kode dibanding library lainnya seperti tensorflow atau numpy.

Scikit-Learn bisa digunakan untuk:

• Classification

• Regression

• Clustering

• Dimensionality Reduction

• Model Selection

• Preprocessing

Anaconda adalah paket distribusi Python dari Continuum Analytics yang berisi paket Python ditambah beberapa paket tambahan untuk keperluan pemrograman data science, matematika hingga teknik dalam satu distribusi platform yang user friendly. Anaconda (Anaconda Distribution) dibuat oleh Contium ini dapat digunakan secara gratis, memiliki banyak package dan tools termasuk Python Distribution di dalamnya. Conda memiliki package management system dan environment system untuk mempermudah pengelolanya. Distribusi ini dapat beroprasi pada sistem operasi Windows, Linux, OS X dan lainnya. Tadinya anaconda dibuat untuk membuat package untuk Python, sekarang ia bisa melakukan itu untuk software lainnya.

1. **Metode Penelitian**

Penelitian ini menggunakan algoritma Naïve Bayes dengan environtment Bahasa pemograman python dan library pendukung sklearn.

## 2.1 Persiapan Data

Data yang akan diuji didapat dari peneleitian sebelumnya berikut:

**Bank Konvensional (file: kon-08-10)**

* X1 = Working capital to total assets (WCTA)
* X2 = Retained earnings to total assets (RETA)
* X3 = Earnings before interest and taxes to total assets(EBITTA)
* X4 = Market value of equity to Total Liabilities (MVETL)
* X5 = Sales to Total Assets

**Bank Syariah (file: bus-10-16)**

* X1 = Non Performing Financing (NFA)
* X2 = Good Corporate Governance (GCG)
* X3 = Return on assets (ROA)
* X4 = Capital Adequacy Ratio (CAR)

## 2.2 Perhitungan Zscore

**Bank Konvensional (file: kon-08-10)**



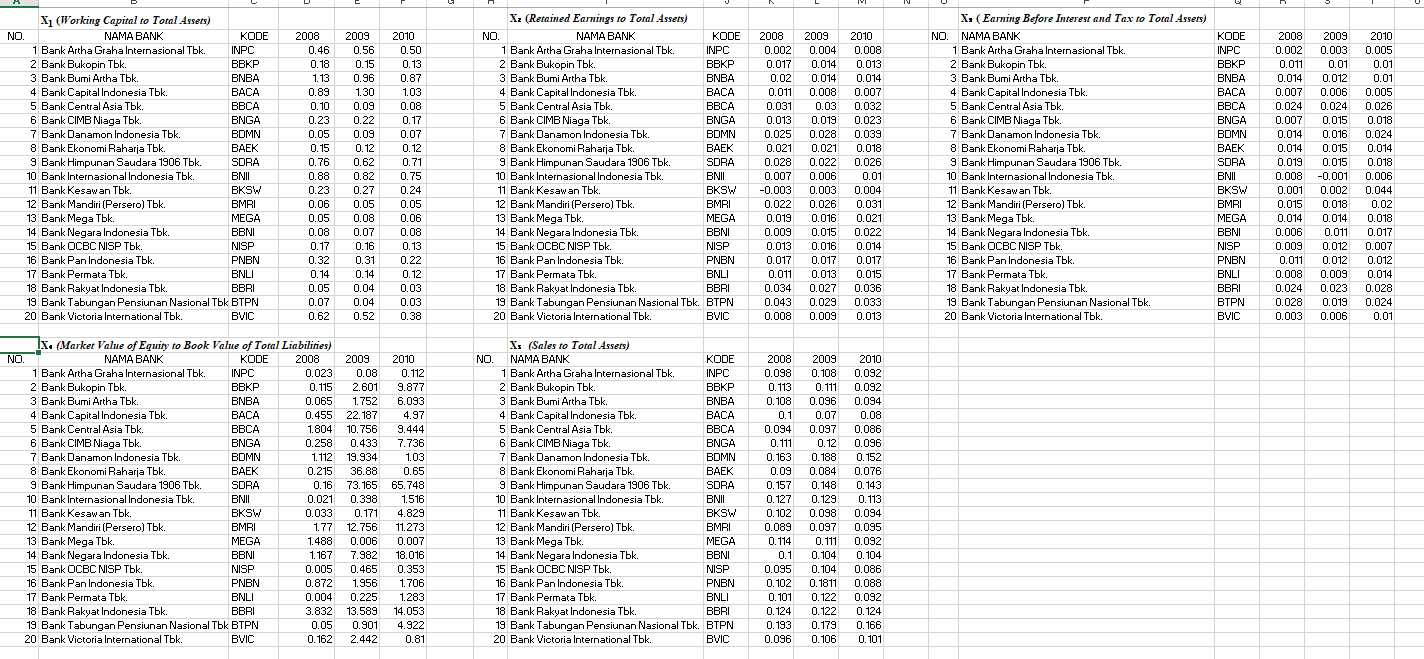
**Bank Syariah (file: bus-10-16)**

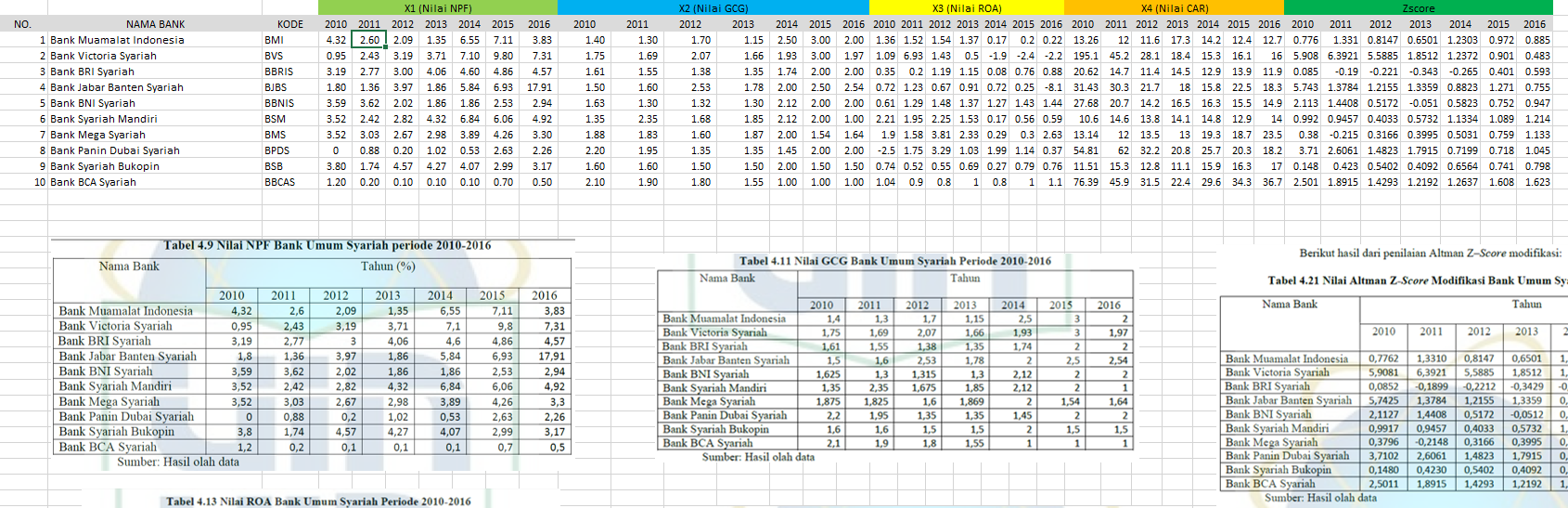


## 2.3 Kategori Class

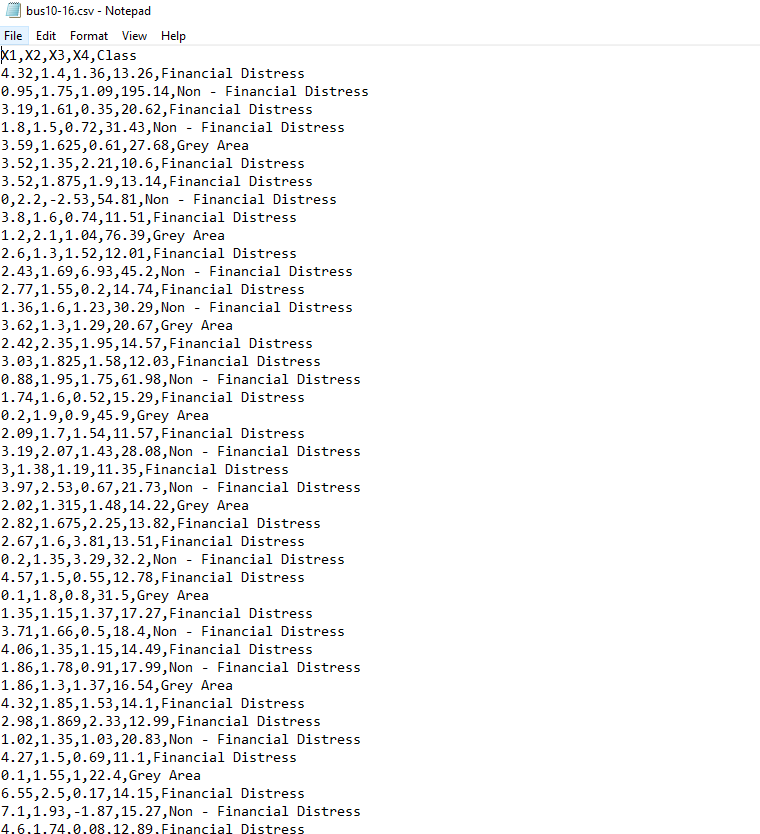
* ZScore < 1.10 = financial distress,
* 1.10 < ZScore < 2.60 = Gray area
* Zscore > 2.60 = non-financial distress

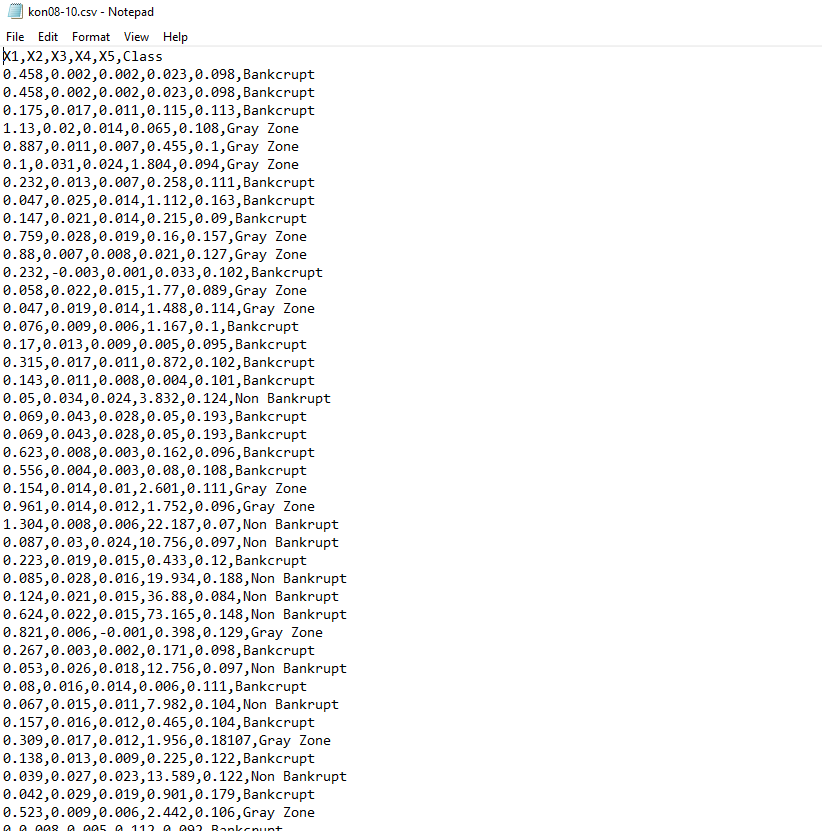
## 2.3 Data Mentah





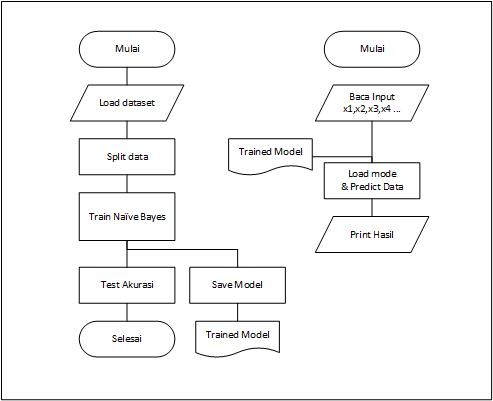
## 2.3 Data Hasil Olah





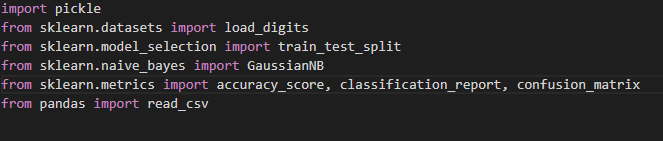
**IMPLEMENTASI DAN PEMBAHASAN**

**3.1 Flowchart Alur**

****

**3.2 Implementasi Pada Python**

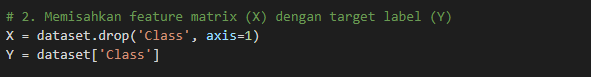
* Import library yang diperlukan untuk pengolahan data seperti berikut



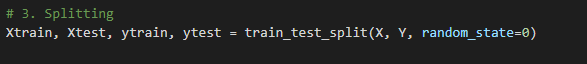
* Script di atas digunakan untuk mengaktifkan package pandas dan sklearn yang akan digunakan pada tahapan analisis. Package pandas sendiri digunakan untuk pengolahan data yang berkaitan dengan data frame. Dan sklearn merupakan library untuk pengaplikasian artificial intelegen pada python



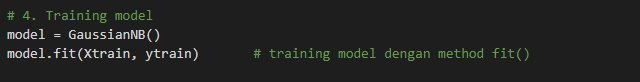
* Baca data dan masukan pada variable dataset dengan perintah read\_csv diikuti nama file



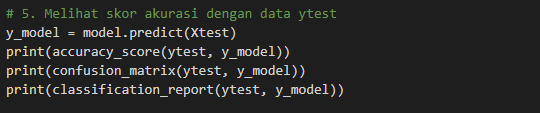
* Pisahkan data baris pertama menjadi label dengan variable X dan sisanya menjadi Y atau data yang akan menjadi data training model



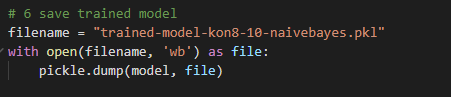
* Split data menjadi data train dan data test akurasi nanti



* Mulai training dengan memanggil algroritma naïve bayes dan masukan kedalam variable lalu jalan perintah fit kepada data training.



* accuracy\_score digunakan untuk menampilkan akurasi mode, classification untuk menampilkan kelas kategori dari model yang telah di uji, dan classification untuk menampilkan informasi akurasi juga rata - rata



* terakhir save model ke dalam sebuah file dengan pickle, supaya model dapat digunakan lagi nanti di program lain.

1. **PENGUJIAN**

**3.1 Hasil Training Model**

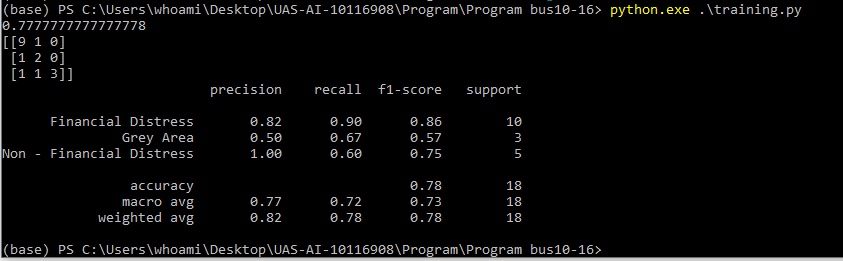


Figure 1 : Akurasi Model Data Bank Syariah

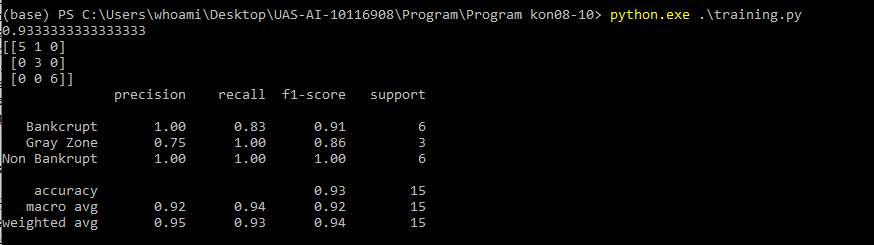
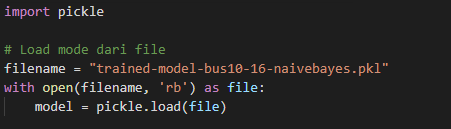


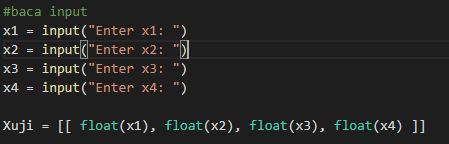
Figure 2 : Akurasi Model Data Bank Konvensional

**3.2 Program Prediksi Berdasarkan Model**

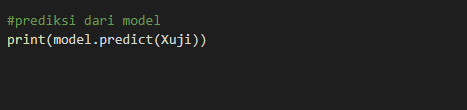
* Pengujian aplikasi dapat dilakukan dengan load model yang telah disimpan sebelumnya ke dalam variabel model



* Baca inputan user dan simpan kedalam array



* Lalu prediksi data tersebut kedalam model dan tampilkan kelayar

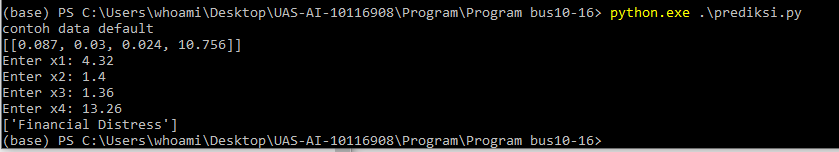


**3.3 Pengetesan Program Model Bank Syariah**

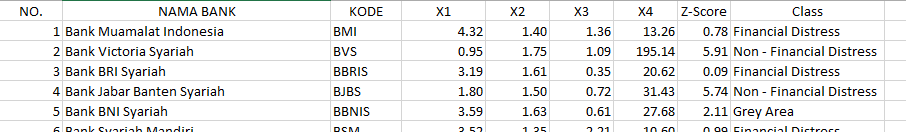
Pengetesan Data Bank Muamalat Indonesia

Input = 4.32, 1.40, 1.36, 13.26

Hasil = Financial Distress



Hasil prediksi sesuai dengan dataset

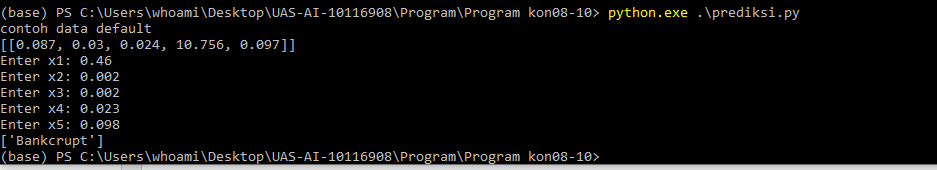


**3.3 Pengetesan Program Model Bank Konvensional**

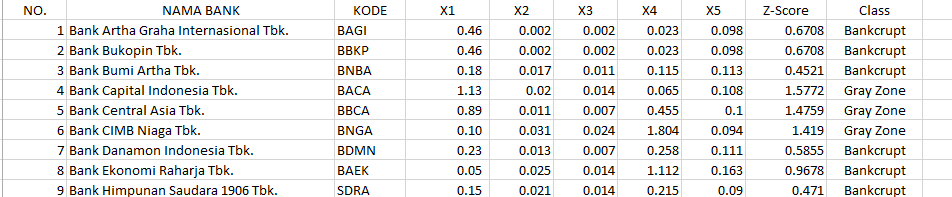
Pengetesan Data Bank Arta Graha Internasional

Input = 0.46, 0.002, 0.002, 0.023, 0.098

Hasil = Bankcrupt



Hasil prediksi sesuai dengan dataset



1. **KESIMPULAN**

Pembuatan model learning menggunakan naïve bayes untuk dataset bank syariah maupun konvensional cukup memuaskan dengan akurasi mencapai 90% pada bank konvensional, dan 80% pada bank syariah. Dimana dengan akurasi tersebut model bisa mendekati hasil asli dataset maupun memprediksi data kedepan.

**DAFTAR PUSTAKA**

[1] M. Adnan Aziz and H. A. Dar, “Predicting corporate bankruptcy: where we stand?” Corporate Governance: The international journal of business in society, vol. 6, no. 1, pp. 18–33, 2006.

[2] Polish companies bankruptcy data Data Set Avaible: http://archive.ics.uci.edu/ml/datasets/Polish+companies+bankruptcy+data [Accessed 01-2-2020]

[3] scikit Learn save model

<https://stackabuse.com/scikit-learn-save-and-restore-models/> [Accessed 01-2-2020]

[4] Machine Learning Algoritma Naïve Bayes,

<https://medium.com/@theinternetbae/membuat-model-machine-learning-dengan-algoritma-naive-bayes-5217cc85df57> [Accessed 01-2-2020]

[5] Introduction to Naïve Bayes Algorithm, Tutorial Points <https://www.tutorialspoint.com/machine_learning_with_python/machine_learning_with_python_classification_algorithms_naive_bayes.htm> [Accessed 01-2-2020]