**Capstone AI-Base Fraude Detection Sytem  
Mendeteksi fraud menggunakan rule-based**

Nama Mentor: Charles Bernando, S.Si., M.A., Ph.D

Logo, company name

Description automatically generated

Disusun oleh:

1. Iyas Abdurahman (3466666/Universitas Pembangunan Nasional Veteran Jakarta)
2. Fiky Ari Wijaya (3108318/Universitas Pancasila )
3. Surya Alfin Maoludin (3390843/Universitas Pendidikan Indonesia)

**Data Analyst and AI**

**Program Studi Independen Angkatan 3**

**PT. MariBelajar Indonesia Cerdas**

**Tahun 2022**

Daftar Isi

[Daftar Isi 2](#_Toc122368809)

[Lembar Pengesahan 3](#_Toc122368810)

[A. Latar Belakang 4](#_Toc122368811)

[B. Perumusan Masalah 4](#_Toc122368812)

[C. Spesifikasi Kebutuhan 5](#_Toc122368813)

[D. Rancangan Solusi 5](#_Toc122368814)

[E. Hasil dan Pembahasan 6](#_Toc122368815)

[F. Kesimpulan 9](#_Toc122368816)

[G. Lampiran 9](#_Toc122368817)

Lembar Pengesahan

**Capstone AI-Base Fraude Detection Sytem  
Mendeteksi fraud menggunakan rule-based**

Disusun oleh:

1. Iyas Abdurahman (3466666/Universitas Pembangunan Nasional Veteran Jakarta)
2. Fiky Ari Wijaya (3108318/Universitas Pancasila )
3. Surya Alfin Maoludin (3390843/Universitas Pendidikan Indonesia)

Disetujui oleh:

Mentor

Charles Bernando, S.Si., M.A., Ph.D

# Latar Belakang

Fraud detection berbasis aturan adalah metode lain yang bergantung pada pola untuk bekerja. Manusia memprogram aturan-aturan ke dalam sistem dan AI. Dalam melakukan pemilahan dalam baik melakukan keputusan dan menyerupai kecerdasan manusia. Maka dibutuhkan nyalah mengoptimisasi parameter dan nilai yang digunakan dalam semua rule sistem untuk mendeteksi fraud berdasarkan Artificial Intelligence (AI).

Maka salah satu hal yang sering terjadi di dunia bisnis ialah tindak fraud, yang dapat merugikan berbagai activitas control, mengingat sebelum adanyan kemajuan mengenai kecerdasan buatan. Segala aktivitas tersebut masih dilakukan secara manual. Yang berarti memungkinkan adanya metode yang kurang tercontrol dari berdampak terhadap kurangnya efektifitas dan efisienensi suatu metode yang berlaku. Mempengaruhi waktu, informasi, kegagalan, ketidak mampuan. sehingga berpengaruh terhadap hasil audit yang kurang sesuai dengan bisnis yang bekerja dengan basis pelanggan yang beragam, global, dan terutama system pengelolaan digital.

Dimana tentu untuk bekerja dalam bisnis, dibutuhkan seperangkat fakta dan aturan untuk memanipulasi suatu data tersebut. Sehingga perlu adanya layanan AI teknologi memilah-milahnya untuk membuat keputusan, menyerupai kecerdasan manusia yang dapat diterapkan sebagai solution menyatakan suatu kondisi. Perlu penyusunan rule base yang berperan pada presisi model, pada tahap pengambilan keputusan ditentukan berdasarkan rancangan yang terbentuk.

# Perumusan Masalah

﻿

Fraud atau biasa disebut dengan kecurangan dalam bidang bisnis, baik yang berasal dari Instansi dalam maupun luar pengelolaan bisnis. Yaitu secara tidak sadar, hal ini muncul karena adanya kebiasaan yang lumrah dan wajar oleh masyarakat umum. Kasus fraud sering terjadi dan bahkan terungkap di setiap tingkatan dan aspek kehidupan masyarakat.

Gejala tindakan fraud terdiri dari ketidak normalan catatan akuntansi, internal control yang rendah, ketidak normalan dalam menganalisis, perubahan gaya hidup, perilaku yang tidak biasa dan tips serta keluhan. Tentang bagaimana segala aktivitas yang masih dilakukan secara manual. Yang berarti memungkinkan adanya metode yang kurang terkontrol dari berdampak terhadap kurangnya efektifitas dan efisienensi suatu metode yang berlaku. maka permasalahan yang akan dikaji pada laporan ini adalah sebagai berikut:

* Bagaimana memahai permasalhan fraud (kecurangan) ?
* Bagaimana peran AI dalam mengoptimalkan layanan detection Fraud ?
* Bagaimana penerapan layanan yang dapat di aplikasikan dalam mengatasi permasalhan seputar Fraude ?

# Spesifikasi Kebutuhan

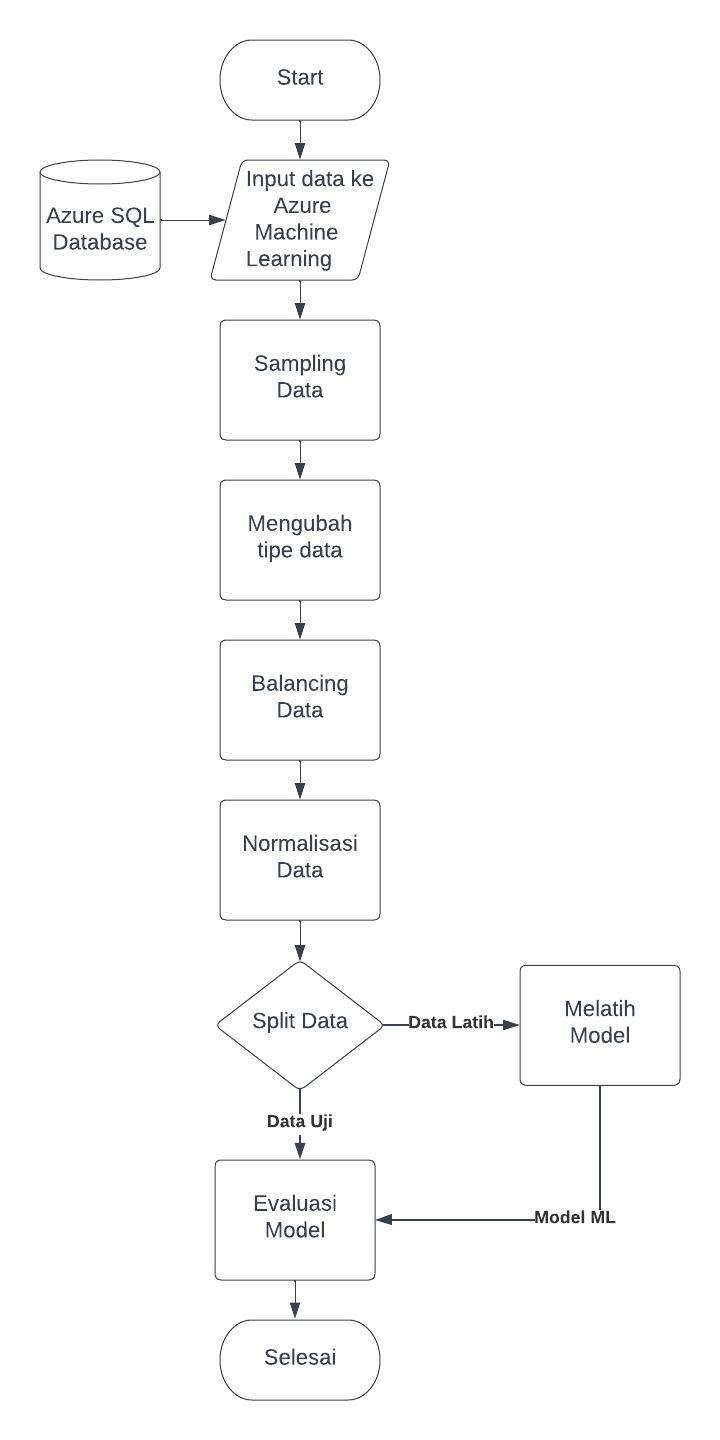
Dalam mengatasi permasalah kecurangan dibutuhkan pengendalian internal baik membangun efektif dan efisien penggunaan yang dapat memudahkan peran pengguna dalam mengaplikasikan nya. Sehingga dalam membangun kebutuhan tersebut pada kajian ini kamu menggunakan beberapa media :

* Microsoft Azure sebagai rangkaian layanan workspace Cloud Computing yang terus berkembang untuk membantu organisasi menghadapi tantangan bisnis.
* Microsoft SQL Server sebagai software jenis Relational Database Management System (RDBMS) dalam mengelola fasilitas backup database sumber data fraud.
* Azure SQL Database adalah layanan cloud pada Microsoft Azure yang memungkinakan pengelolaan entitas layaknya SQL server database.
* Azure Machine Learning Designer layanan cloud pada Microsoft Azure yang memungkinakan penyediaan ruang yang digunakan untuk melatih dan mengelola model pembelajaran mesin, dan diperlukan sumber daya.
* Cognitive Service, layanan yang menyediakan berbagai kebutuhan yang dibutuhkan oleh pengembang untuk membuat sebuah sistem dengan kemampuan kecerdasan buatan yang terdapat pada cloud Microsoft Azure.

# Rancangan Solusi

Rancangan solusi yang kami kembangkan akan melibatkan pendekatan machine learning dalam penerapannya. Kami akan menggunakan model machine learning untuk menentukan apakah transaksi tersebut merupakan transaksi fraud atau tidak. Kami juga akan memanfaatkan cloud dalam penerapannya untuk mengurangi kerumitan dalam membuat model machine learningnya. Untuk itu, kami akan menggunakan beberapa tools yang ada di cloud seperti Azure SQL Database sebagai media penyimpanan, dan Azure Machine Learning untuk pengembangan model machine learning.

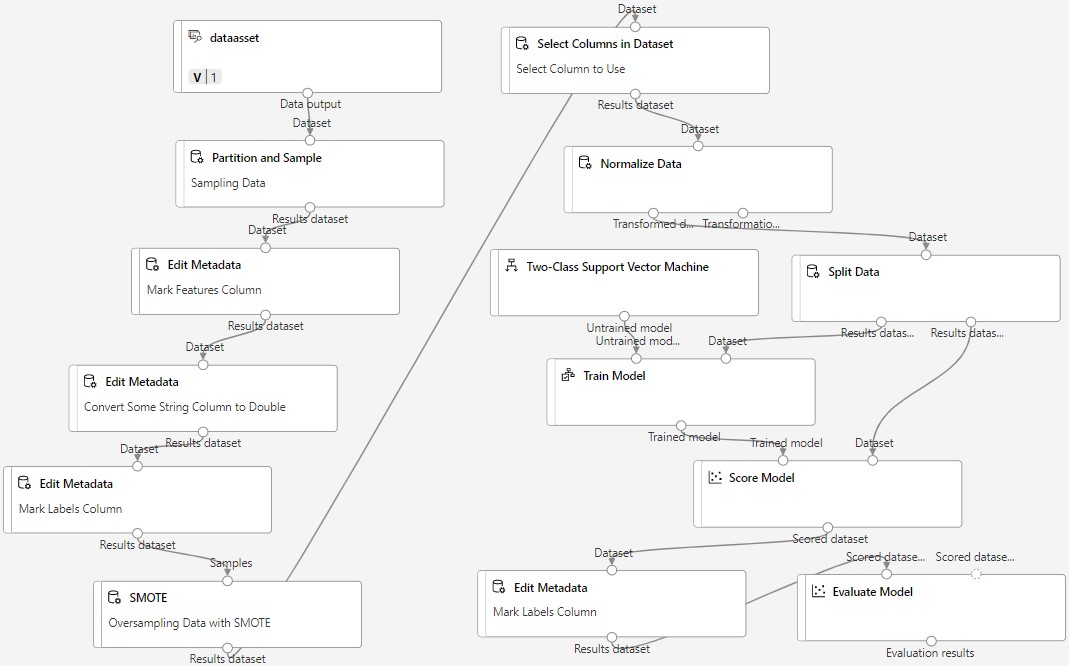
Berikut adalah flowchart rancangan solusi yang telah kami susun:



# Hasil dan Pembahasan

Untuk solusi machine learning yang kami susun dalam Capstone Project ini akan digambarkan lebih lanjut dalam pipeline machine learning.

Berikut adalah pipeline machine learning yang kami buat:



Pipeline dimulai dari komponen dataasset yang merupakan dataset yang telah diambil dari azure sql database dan kemudian di kapsulasi menjadi komponen dataasset pada azure machine learning.

Setelah itu, akan dilakukan data sampling untuk mengambil 1/100 data secara random dari dataasset. Hal ini dilakukan karena dataasset mempunyai terlalu banyak data (sekitar 6,000,000 baris) dan hal tersebut dapat membuat beban komputasi yang dilakukan oleh cloud terlalu besar sehingga berbagai tahapan machine learning yang akan dilakukan nantinya akan diproses dengan lebih lama.

Lalu, akan dilakukan beberapa tahapan edit metadata, seperti mark feature column untuk menandai kolom yang merupakan kolom fitur, convert some string column to double untuk mengconvert beberapa kolom string menjadi double, dan mark labels column untuk menandai kolom yang merupakan kolom label.

Setelah itu akan dilakukan tahapan oversampling data. Tahapan ini merupakan salah satu tahapan kritis pada pipeline ini karena permasalahan fraud itu sendiri merupakan salah satu jenis permasalahan yang datanya sangat tidak seimbang antara data berlabel fraud dan data berlabel bukan fraud. Hal ini tentunya merupakan tantangan tersendiri bagi kami untuk mengatasi hal tersebut, dan tentunya kami atasi dengan melakukan oversampling data SMOTE.

Tahapan selanjutnya adalah memilih kolom mana saja yang akan digunakan, karena tentunya pasti ada beberapa kolom yang jika digunakan pun tidak akan mempengaruhi apa-apa seperti kolom nama, alamat, kode pos, dll. Tahapan ini juga dapat mengurangi beban komputasi pada cloud saat memproses karena jumlah kolom yang digunakan berkurang.

Tahapan selanjutnya adalah normalisasi data, yaitu menyamakan range data bertipe angka untuk mengurangi bias antara setiap kolom yang dapat terjadi jika range data setiap kolom berbeda.

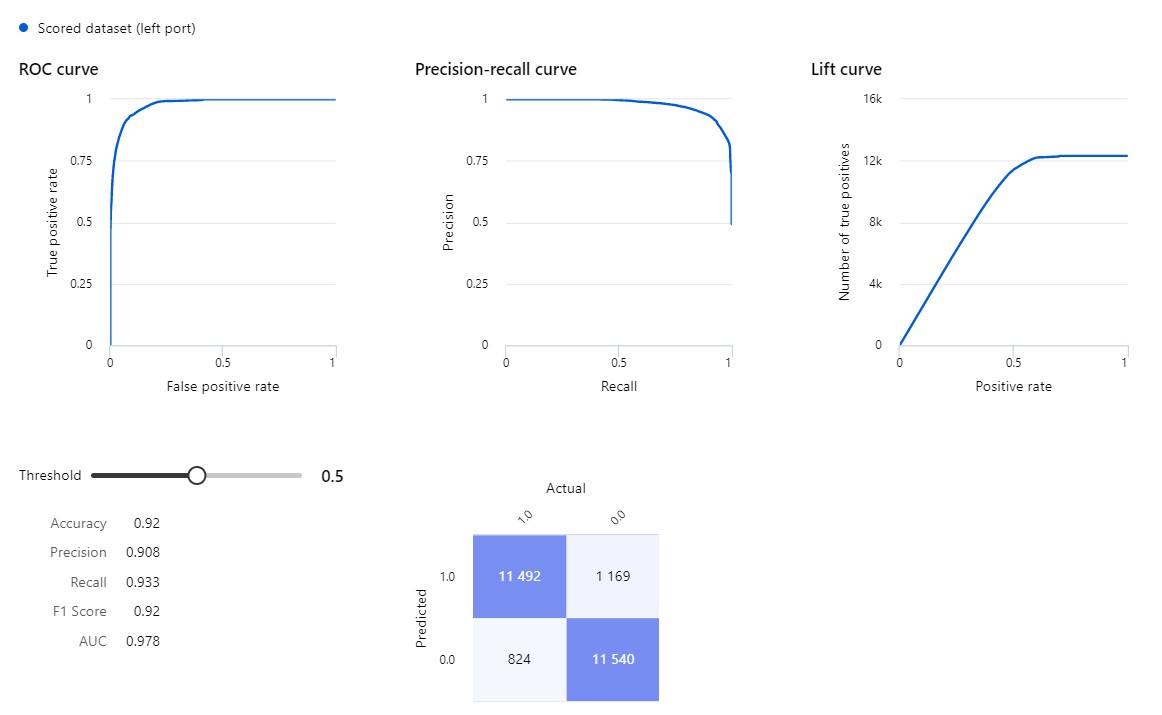
Kemudian dilakukan tahapan split data dengan rasio 80% data latih dan 20% data testing.

Setelah itu, dilakukan tahapan train model dengan Support Vector Machine untuk melatih model yang nantinya akan digunakan untuk menentukan apakah data tersebut merupakan fraud atau bukan.

Lalu, dilakukan tahapan score model untuk menguji model terhadap inputan data testing, label apa yang akan diberikan model terhadap setiap data testing apakah akan sama dengan label asli yang ada pada data testing.

Lalu, sebelum dilakukan tahapan evaluasi model akan dilakukan tahapan edit metadata terlebih dahulu, yaitu menandai kolom yang merupakan kolom label.

Terakhir dilakukan tahapan evaluasi model, yang hasilnya dapat dilihat pada gambar dibawah ini:



Dari gambar diatas, dapat dilihat bahwa evaluasi matriks yang dihasilkan cukup memuaskan dengan akurasi, presisi, recall, dan f1-score semuanya dengan nilai diatas 0.9.

Batasan dari solusi yang kami kembangkan ini mungkin adalah karena kami melakukan sampling 1/100 pada dataset, mungkin hasil evaluasi yang ada pada gambar diatas kurang dapat melambangkan hasil evaluasi dari keseluruhan dataset karena data yang digunakan hanya 1/100 dari keseluruhan dataset.

# Kesimpulan

* Sebuah kecurangan di dalam bisnis bisa kita deteksi dengan menggunakan AI.
* Gejala tindakan fraud terdiri dari ketidak normalan catatan akuntansi, internal control yang rendah, ketidak normalan dalam menganalisis, perubahan gaya hidup, perilaku yang tidak biasa dan tips serta keluhan.
* Rancangan solusi yang kami sajikan terhadap permasalahan kecurangan (fraud) dalam bisnis ini adalah membuat sebuah layanan pendeteksi kecurangan dengan menggunakan kecerdasan buatan (AI) yang kami namakan “***AI based Fraud Detection System***”
* Untuk membuat sebuah service AI diperlukan beberapa layanan diantaranya yaitu Microsoft Azure, Microsoft SQL Server, Azure SQL Database, Azure Machine Learning Designer, Azure Cognitive Service.
* Alur kerja dari rancangan yang kami buat dimulai dari pemasukan dataset kedalam Azure SQL Database; dilanjut dengan Input data ke Azure Machine Learning; kemudian dilakukan sampling data; Pengubahan tipe data; Balancing Data; Normalisasi data; Lalu Split data untuk melakukan Training model; Terakhir kita melakukan evaluasi terhadap model yang dibuat.

# Lampiran

URL Github: <https://github.com/Alfinnnnn/Capstone_kel17>

Link Youtube pitching video: <https://youtu.be/nZ2R7Outk5w>

Link Youtube Demonstrasi video: <https://youtu.be/_pD6I8LatfQ>