

**LAPORAN AKHIR**

**STUDI INDEPENDEN BERSERTIFIKAT**

**Artificial Intelligence for Gen Y**

**PT Orbit Ventura Indonesia**

*Diajukan untuk memenuhi persyaratan kelulusan Program Magang dan Studi Independen Bersertifikat Kampus Merdeka*



Disusun oleh :

**MUHAMAD RENDI**

**5190411597**

**PROGRAM STUDI INFORMATIKA**

**FAKULTAS SAINS & TEKNOLOOGI**

**UNIVERSITAS TEKNOLOGI YOGYAKARTA**

**2022**

**Lembar Pengesahan**

Program Studi Informatika – Universitas Teknologi Yogyakarta

Studi Independen Bersertifikat Kampus Merdeka

Artificial Intelligence for Gen Y

Di PT Orbit Ventura Indonesia

Disusun oleh :

**Muhamad Rendi / 5190411597**

Disetujui dan disahkan sebagai

Laporan Magang atau Studi Independen Bersertifikat Kampus Merdeka

Yogyakarta, 28 Januari 2022

Ketua Program Studi Informatika

Universitas Teknologi Yogyakarta

**Dr. Enny Itje Sella, S.Si., M.Kom.**

NIK. 111116086

# LEMBAR PENGESAHAN

Studi Independen Bersertifikat Kampus Merdeka

Artificial Intelligence for Gen Y

Di PT Orbit Ventura Indonesia

Disusun oleh :

**Muhamad Rendi / 5190411597**

disetujui dan disahkan sebagai

Laporan Magang atau Studi Independen Bersertifikat Kampus Merdeka

Jakarta, 28 Januari 2022

Mentor Kegiatan

PT Orbit Ventura Indonesia



Dino Febriyanto, S. Kom

NIK. 20202042

# ABSTRAK

Program Magang dan Studi Independen Bersertifikat (MSIB) yang merupakan salah satu metode pembelajaran yang dicanangkan dalam kebijakan Kampus Merdeka. PT Orbit Ventura Indonesia merupakan salah satu mitra yang bergabung dalam program tersebut. Salah satu program yang ditawarkan oleh PT Orbit Ventura Indonesia yaitu Artificial Intelligence for Gen Y yang berdurasi 5 bulan dan salah satu persyaratan kelulusan program ini adalah membuat proyek akhir yang memodelkan *artificial intelligence* menjadi suatu produk aplikasi sebagai bentuk implementasi pengetahuan yang telah didapatkan oleh penulis selama proses pembelajaran di PT. Orbit Ventura Indonesia. Penulis membuat proyek akhir “Prediksi Kemungkinan Sesorang Dapat Menderita Stroke menggunakan algoritma K-Nearest Neighbor (K-NN) dan Multi-layer Perceptron” yang merupakan solusi dari isu kesehatan menggunakan salah domain *artificial intelligence.* Kemudian untuk pengujian performanya dilakukan dengan menggunakan matrikx akurasi, f1-score, recall dan precision. Setelah dilakukan pengujian, model yang dirancang menggunakan K-NN menghasikan skor akurasi 89%, precision 90%, recall 89%, dan f1-score 89%. Sedangkan model yang dirancang menggunakan Multi-layer Perceptron menghasilkan skor akurasi 88%, precision 88%, recall 88%, dan f1-score 88%. Skor yang ditunjukkan menandakan bahwa hasil prediksi dari model yang dirancang dengan algoritma K-NN memiliki hasil yang lebih baik dari pada model yang dirancang dengan algoritma Multi-layer Perceptron.

# KATA PENGANTAR

Puji syukur kepada Tuhan Yang Maha Esa, karena dengan rahmat dan karunia-Nya, penulis dapat menyelesaikan Laporan Akhir Magang dan Studi Independen Bersertifikat yang berjudul “Studi Independen Bersertifikat AI for Gen Y di PT. Orbit Ventura Indonesia”. Adapun maksud dari laporan ini adalah sebagai salah satu syarat untuk menyelesaikan Program Magang dan Studi Independen Bersertifikat (MSIB) Kampus Merdeka.

Penulis mengucapkan terima kasih sebesar – besarnya kepada:

1. Dosen-dosen Universitas Bengkulu yang senantiasa membimbing saya.
2. PT. Orbit Ventura Indonesia yang telah membuka peluang untuk mempelajari *Aritificial Intelligence* melalui program Magang dan Studi Independen Bersertifikat (MSIB) Kampus Merdeka
3. Orang tua saya tercinta yang selalu mendukung saya.
4. Teman-teman saya yang selalu memberi semangat kepada saya.
5. Akhir kata, penulis berharap semoga Laporan Akhir ini dapat bermanfaat bagi semua pihak yang membacanya.

Yogyakarta, 28 Januari 2022

Penulis

# DAFTAR ISI

[LEMBAR PENGESAHAN iii](#_Toc94474604)

[ABSTRAK iv](#_Toc94474605)

[KATA PENGANTAR v](#_Toc94474606)

[DAFTAR ISI vi](#_Toc94474607)

[BAB I 1](#_Toc94474608)

[1.1. Latar belakang 1](#_Toc94474609)

[1.2. Lingkup 1](#_Toc94474610)

[1.3. Tujuan 2](#_Toc94474611)

[BAB II GAMBARAN UMUM PT. ORBIT VENTURA INDONESIA 3](#_Toc94474612)

[2.1. Struktur Organisasi 3](#_Toc94474613)

[2.2. Lingkup Pekerjaan 3](#_Toc94474614)

[2.3. Deskripsi Pekerjaan 3](#_Toc94474615)

[2.4. Jadwal Kerja 4](#_Toc94474616)

[BAB III 5](#_Toc94474617)

[3.1. Deskripsi Proyek 5](#_Toc94474618)

[3.2. AI Project Cycle 5](#_Toc94474619)

[a. Problem Scoping 6](#_Toc94474620)

[b. Data Acquisition 6](#_Toc94474621)

[c. Data Exploration 7](#_Toc94474622)

[d. Modelling 7](#_Toc94474623)

[e. Evaluation 8](#_Toc94474624)

[3.3. Hambatan, Solusi, Hasil, dan Capaian 8](#_Toc94474625)

[BAB IV 11](#_Toc94474626)

[4.1 Kesimpulan 11](#_Toc94474627)

[4.2 Saran 11](#_Toc94474628)

[REFERENSI 12](#_Toc94474629)

[LAMPIRAN 20](#_Toc94474630)

[A. TOR 20](#_Toc94474631)

[B. LOG ACTIVITY 21](#_Toc94474632)

[C. Dokumen Teknik 25](#_Toc94474633)

[● AI Project Cycle 25](#_Toc94474634)

[● Profil Peneliti 26](#_Toc94474635)

# 

# BAB I

**PENDAHULUAN**

## Latar belakang

Program Magang dan Studi Independen Bersertifikat (MSIB) yang merupakan salah satu metode pembelajaran yang dicanangkan dalam kebijakan Kampus Merdeka. Kebijakan ini diluncurkan dalam rangka menyiapkan mahasiswa-mahasiswa Indonesia menghadapi perubahan sosial, budaya, dunia kerja dan kemajuan teknologi yang pesat, kompetensi mahasiswa harus disiapkan untuk lebih gayut dengan kebutuhan zaman.

PT Orbit Ventura Indonesia merupakan salah satu mitra yang bergabung dalam program Studi Independen Bersertifikat Kampus Merdeka. PT Orbit Ventura Indonesia telah didirikan pada tahun 2015 untuk memfasilitasi pemahaman yang lebih baik tentang bagaimana ilmu pengetahuan dan teknologi dapat meningkatkan kehidupan negara berkembang. Salah satu program yang ditawarkan oleh PT Orbit Ventura Indonesia dalam Magang dan Studi Independen Bersertifikat Kampus Merdeka yaitu Artificial Intelligence for Gen Y yang diperuntukkan untuk para pencari kerja. Orbit AI for Gen Y adalah program pelatihan *Artificial Intelligence* holistik yang *hybrid* dan inklusif untuk siswa dari semua sektor dan minat yang beragam. Program ini menggunakan pendekatan *Deep Learning* dan berbasis proyek. Pembelajaran dalam *AI for Gen Y* mencakup mulai dari level dasar pengenalan AI, hingga tingkat lanjut yang mencakup pemrograman *Python* dan *AI Project Cycle (Problem Scoping, Data Acquisition, Data Exploration, Modelling, Evaluation & Deployment)*. Selain itu program ini juga dilengkapi dengan kursus *Life Skills* yang penting dan diperlukan untuk kesuksesan karir.

Program *AI for Gen Y* berdurasi 5 bulan dan salah satu persyaratan kelulusan program ini adalah membuat proyek akhir yang memodelkan *artificial intelligence* menjadi suatu produk aplikasi sebagai bentuk implementasi pengetahuan yang telah didapatkan oleh penulis selama proses pembelajaran di PT. Orbit Ventura Indonesia.

## Lingkup

Proyek akhir berisi tentang bagaimana tahapan-tahapan atau memodelkan sistem *artificial intelligence* menjadi suatu produk aplikasi. Penulis membuat *paper* berjudul “Prediksi Kemungkinan Sesorang Dapat Menderita Stroke menggunakan algoritma K-Nearest Neighbor (K-NN) dan Multi-layer Perceptron” yang mengaplikasikan salah satu domain *artificial intelligence*.

## Tujuan

Tujuan yang ingin dicapai dalam Studi Independen Bersertifikat Kampus Merdeka adalah sebagai berikut.

1. Mewujudkan gagasan mahasiswa dalam mengembangkan produk inovatif yang menjadi gagasannya, dalam hal ini mahasiswa diharapkan dapat membuat solusi dari isu sosial menggunakan domain *artificial intelligence*.
2. Menyelenggarakan pendidikan berbasis riset dan pengembangan.
3. Meningkatkan prestasi mahasiswa dalam ajang nasional dan internasional.

Sedangkan tujuan yang ingin dicapai dengan dibuatnya proyek akhir berjudul “Prediksi Kemungkinan Sesorang Dapat Menderita Stroke menggunakan algoritma K-Nearest Neighbor (K-NN) dan Multi-layer Perceptron” adalah sebagai berikut.

1. Mengimplementasikan ilmu yang telah didapat oleh penulis selama mengikuti program Studi Independen Bersertifikat dalam bentuk model Machine Learning yang dapat digunakan dan dikembangkan dalam berbagai platform untuk kesejahteraan masyarakat dalam bidang kesehatan.
2. Membantu masyarakat dan para tenaga kesehatan dalam mengenali orang-orang yang beresiko terkena penyakit stroke.

# BAB II GAMBARAN UMUM PT. ORBIT VENTURA INDONESIA

## Struktur Organisasi

Orbit merupakan warisan peninggalan Almarhum Dr. Hasri Ainun Habibie, Ibu Negara RI dan istri Almarhum Prof. Dr. Ing. B J Habibie – Presiden ke 3 Republik Indonesia. Beliau merupakan sosok penggerak perkembangan dunia Pendidikan di Indonesia.

Visi: PT Orbit Ventura Indonesia didirikan pada tahun 2016 dengan tujuan untuk meningkatkan kualitas hidup melalui inovasi, edukasi, dan pelatihan keterampilan.

*Untuk memberikan pembelajaran berbasis keterampilan transformatif terbaik, untuk pencari kerja & pencipta kerja.*

Misi:

* Untuk membangun jaringan Orbit Transformation Center (OTC) secara nasional untuk menyampaikan kurikulum keterampilan masa depan berbasis sertifikasi melalui Platform Konten Digital.
* Untuk secara proaktif bekerja dengan pemerintah & organisasi dengan mengubah tenaga kerja mereka agar sesuai dengan perubahan pekerjaan yang terjadi karena Industri 4.0.
* Untuk melatih pemuda dengan keterampilan kewirausahaan & mencocokkan mereka dengan peluang masa depan yang muncul di berbagai industri.
* Untuk menghubungkan jaringan inkubator dan akselerator yang dikurasi ke industry, investor, dan ekosistem startup global.

## Lingkup Pekerjaan

Program ini dapat diikuti dari mana saja di Indonesia terlepas dari masalah dan tantangan pandemi. Program ini telah dirancang untuk pandemi sesuai dengan “Orbit Way" kami untuk hasil akhir yang luar biasa. "Orbit Way" menggunakan model menarik yang memadukan sesi kursus online yang dipimpin instruktur, dengan program mentor secara live, dan penilaian proyek akhir tatap langsung.

## Deskripsi Pekerjaan

Pada program Artificial Intelligence for Gen Y, dilakukan kegiatan pembelajaran terkait *artificial intelligence* (AI) dengan luaran berupa proyek akhir yang berbentuk aplikasi atau program dengan penerapan AI di dalamnya. Penulis melaksanakan proyek akhir “Prediksi Kemungkinan Sesorang Dapat Menderita Stroke menggunakan algoritma K-Nearest Neighbor (K-NN) dan Multi-layer Perceptron” sesuai dengan *AI Project Cycle,* yakni *Problem Scoping, Data Acquisition, Data Exploration, Modelling,* hingga *Evaluation.* Program ini dapat diikuti dari mana saja di Indonesia, program ini memiliki jalur langsung menuju peluang kerja dan memenuhi aspirasi kewirausahaan.

## Jadwal Kerja

Program ini berlangsung setiap hari Senin sampai dengan Jumat selama 3 jam dalam satu hari nya. Program berlangsung sejak Agustus 2021 sampai Februari 2022.

Program pagi dari jam 08.00 WIB - 11.15 WIB dan

Program sore dari jam 15.30 WIB - 18.15 WIB.

# BAB III

**GAMBARAN UMUM PROYEK AKHIR**

**“****Prediksi Kemungkinan Sesorang Dapat Menderita Stroke menggunakan algoritma K-Nearest Neighbor (K-NN) dan Multi-layer Perceptron”**

## Deskripsi Proyek

Menurut WHO (The World Health Organization) definisi stroke adalah berkembang pesat tanda-tanda klinis gangguan fungsi otak fokal (atau global), dengan gejala berlangsung 24 jam atau lebih atau menyebabkan kematian, tanpa penyebab yang jelas selain dari vaskular[1]. Stroke adalah kondisi yang terjadi ketika pasokan darah ke otak terganggu atau berkurang akibat penyumbatan (stroke iskemik) atau pecahnya pembuluh darah (stroke hemoragik). Tanpa darah, otak tidak akan mendapatkan asupan oksigen dan nutrisi, sehingga sel-sel pada sebagian area otak akan mati. Kondisi ini menyebabkan bagian tubuh yang dikendalikan oleh area otak yang rusak tidak dapat berfungsi dengan baik. Faktor yang menyebabkan terjadinya stroke antara lain adalah faktor kesehatan, faktor gaya hidup, faktor usia bahkan faktor keturunan.

Diperkirakan 17,9 juta orang meninggal karena CVD (Penyakit kardiovaskular) pada 2019, mewakili 32% dari semua kematian global. Dari kematian tersebut, 85% disebabkan oleh serangan jantung dan stroke menurut laporan dari WHO (The World Health Organization)[2]. Berdasarkan laporan dari Pusat Data dan Informasi Kementrian Kesehatan RI, secara nasional prevalensi stroke di Indonesia tahun 2018 berdasarkan diagnosis dokter pada penduduk berumur diatas 15 tahun sebesar 10,9%, atau diperkirakan sebanyak 2.120.362 orang[3 ]. Maka dari itu diperlukan kesadaran bagi semua orang untuk menghindari resiko-resiko yang dapat menyebabkan penyakit stroke.

Berdasarkan pernyataan diatas muncul ide peneliti untuk membuat model machine learning tentang Prediksi Kemungkinan Sesorang Dapat Menderita Stroke menggunakan algoritma K-Nearest Neighbor (K-NN) dan Multi-layer Perceptron yang dimana tujuan dari proyek ini yaitu untuk mengklasifikasikan orang-orang yang beresiko terkena penyakit stroke atau tidak. Dengan adanya model ini diharapkan semua orang dapat mampu menjaga kesehatan agar tidak beresiko terkena penyakit stroke dengan melakukan gaya hidup sehat dan selalu menjaga kesehatan tubuh. Implementasinya model ini dapat dijalankan pada sebuah platform aplikasi web ataupun android.

## AI Project Cycle

### Problem Scoping

Pendefinisian ruang lingkup masalah yang akan diselesaikan dilakukan dengan pendekatan 4W (*Who, What, Where, Why*). Berikut adalah *problem scoping* sebelum dilakukannya pengembangan aplikasi “Prediksi Kemungkinan Sesorang Dapat Menderita Stroke menggunakan algoritma K-Nearest Neighbor (K-NN) dan Multi-layer Perceptron”.

***Who* - Siapa yang memiliki masalah?**

Masyarakat yang memiliki riwayat stroke dan masyarakat indonesia

***What* - Apa permasalahan sebenarnya?**

Berdasarkan laporan dari Pusat Data dan Informasi Kementrian Kesehatan RI, secara nasional prevalensi stroke di Indonesia tahun 2018 berdasarkan diagnosis dokter pada penduduk berumur diatas 15 tahun sebesar 10,9%, atau diperkirakan sebanyak 2.120.362 orang[3]. Maka dari itu diperlukan kesadaran bagi semua orang untuk menghindari resiko-resiko yang dapat menyebabkan penyakit stroke. Salah satunya pada proyek ini, dimana akan dibuat sebuah model machine learning untuk mengklasifikasikan orang-orang yang beresiko terkena penyakit stroke atau tidak.

***Where* - Dimana/pada saat apa permasalahan ini muncul?**

Pada saat masyarakat mulai beranjak menua dan memiliki gaya hidup yang tidak sehat, dimana hal tersebut dapat menimbulkan resiko terkena serangan jantung

***Why* - Mengapa masalah ini sangat penting untuk dibahas?**

Karena menurut laporan dari WHO (The World Health Organization) Diperkirakan 17,9 juta orang meninggal karena CVD (Penyakit kardiovaskular) pada 2019, mewakili 32% dari semua kematian global. Dari kematian tersebut, 85% disebabkan oleh serangan jantung dan stroke. Hal ini menandakan bahwa panyakit stroke dapat dikategori berbahaya untuk keselamatan hidup seseorang. Maka dari itu diperlukan kesadaran bagi semua orang untuk menghindari resiko-resiko yang dapat menyebabkan penyakit stroke.

### Data Acquisition

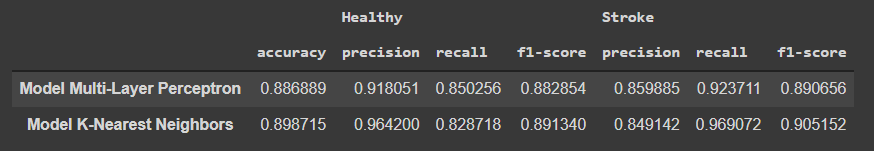
Pada penelitian “Prediksi Kemungkinan Sesorang Dapat Menderita Stroke menggunakan algoritma K-Nearest Neighbor (K-NN) dan Multi-layer Perceptron” ini, dibutuhkan data berupa *dataset stroke* yang diperoleh melalui situs kaggle (<https://www.kaggle.com/fedesoriano/stroke-prediction-dataset>). Kumpulan dataset tersebut akan digunakan untuk melatih dan mengembangkan model AI yang dapat memprediksikan Kemungkinan Sesorang Dapat Menderita Stroke

### Data Exploration

Data yang telah dikumpulkan melalui tahap *data acquisition* belum sepenuhnya bersih dan siap untuk digunakan, sehingga perlu melewati tahap *preprocessing* terlebih dahulu. Setelah dilakukan eksplorasi, ternyata ada banyak *noise* dan *imbalance* pada dataset yang telah didapatkan, seperti *missing data*, *outliers*, dan lain sebagainya. Untuk membersihkan dataset dari *noise* yang ada, kami melakukan analisis terlebih dahulu menggunakan beberapa library python seperti pandas, matplotlib dan lain sebagainya. Setelah dilakukan analisis, selanjutnya kami melakukan perbaikan struktur datanya menggunakan code dengan python. Setelah melakukan pembersihan data, selanjutnya kami melakukan *attribute selection* untuk memilih atribut-atribut yang akan digunakan pada dataset yang ada.

### Modelling

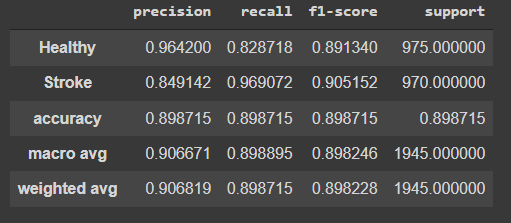
Dalam mengklasifikasi Kemungkinan Sesorang Dapat Menderita Stroke mengggunakan data yang ada, peneliti menggunakan K-Nearest Neighbor dan Multi-layer Perceptron. Pemilihan algoritma ini didasarkan pada dataset yang dimiliki peneliti memiliki data berkategori dan data numerik sehingga cocok menggunakan 2 algoritma tersebut. Implementasi algoritma tersebut dilakukan melalui Google Colab dengan library dari scikit-learn. Model yang dibuat berupa klasifikasi dengan 2 algoritma yang digunakan peneliti.



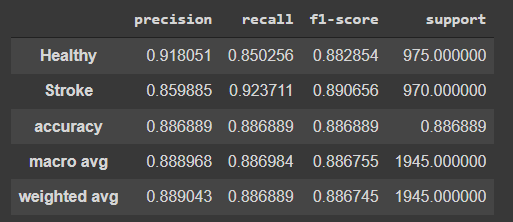
Gambar 3.1 Hasil modeling 2 algortima

### Evaluation

Untuk pengujian performanya dilakukan dengan menggunakan matrikx akurasi, f1-score, recall dan precision. Setelah dilakukan pengujian, model yang dirancang menggunakan K-NN menghasikan skor akurasi 89%, precision 90%, recall 89%, dan f1-score 89%. Sedangkan model yang dirancang menggunakan Multi-layer Perceptron menghasilkan skor akurasi 88%, precision 88%, recall 88%, dan f1-score 88%. Skor yang ditunjukkan menandakan bahwa hasil prediksi dari model yang dirancang dengan algoritma K-NN memiliki hasil yang lebih baik dari pada model yang dirancang dengan algoritma Multi-layer Perceptron.



Gambar 3.2. Performas Model K-Nearest Neighboar



Gambar 3.3. Performa model Multi-layer Perceptron

## Hambatan, Solusi, Hasil, dan Capaian

Selama pengerjaan proyek akhir ini, terdapat beberapa hal yang menghambat progres proyek akhir hingga akhirnya ditemukan solusi dengan rincian sebagai berikut.

Tabel Hambatan Pengerjaan Proyek Akhir beserta Solusi

|  |  |  |
| --- | --- | --- |
| **No.** | **Hambatan** | **Solusi** |
| 1. | Kolaborasi pengerjaan proyek yang harus dilaksanakan secara daring sepenuhnya. | Mencari *tools* yang mendukung pengerjaan proyek secara *online* dan *collaborative*, meliputi Google Colab, Visual Studio Code. |
| 2. | Peneliti memiliki kesibukan tersendiri sehingga sulit untuk menyelesaikan project | Menetepkan waktu untuk menyelesaikan project ini sampai tuntas yaitu setiap malam jam 22.00 WIB |

Sementara itu, luaran yang dicapai dari proyek akhir “Prediksi Kemungkinan Sesorang Dapat Menderita Stroke menggunakan algoritma K-Nearest Neighbor (K-NN) dan Multi-layer Perceptron” adalah mencakup berikut ini, dimana luaran tersebut lebih lanjut dijelaskan pada Lampiran C.

* + 1. Implementasimodel *machine learning* menggunakan bahasa pemrograman Python.
    2. Dokumen teknis AI *project life cycle* yang berisi ringkasan dari tahapan pengembangan Prediksi Kemungkinan Sesorang Dapat Menderita Stroke menggunakan algoritma K-Nearest Neighbor (K-NN) dan Multi-layer Perceptron, mulai dari *problem scoping* hingga *Evaluation*.

Dengan demikian, dari proyek akhir “Prediksi Kemungkinan Sesorang Dapat Menderita Stroke menggunakan algoritma K-Nearest Neighbor (K-NN) dan Multi-layer Perceptron”, telah dicapai tujuan di bawah ini sebagaimana disampaikan sebelumnya pada sub bab I.3.

1. Ketercapaian tujuan Studi Independen Bersertifikat Kampus Merdeka:
   1. Terwujudnya produk hasil gagasan inovatif “Prediksi Kemungkinan Sesorang Dapat Menderita Stroke menggunakan algoritma K-Nearest Neighbor (K-NN) dan Multi-layer Perceptron” yang ditujukan untuk membantu masyarakat dan para tenaga kesehatan dalam mengenali orang-orang yang beresiko terkena penyakit stroke
   2. Terselenggaranya pendidikan AI for Gen Y berbasis riset dan pengembangan yang dibuktikan dengan terselesaikannya proyek akhir “Prediksi Kemungkinan Sesorang Dapat Menderita Stroke menggunakan algoritma K-Nearest Neighbor (K-NN) dan Multi-layer Perceptron” ini.
2. Ketercapaian tujuan proyek akhir “Prediksi Kemungkinan Sesorang Dapat Menderita Stroke menggunakan algoritma K-Nearest Neighbor (K-NN) dan Multi-layer Perceptron”:
3. Terciptanya “Prediksi Kemungkinan Sesorang Dapat Menderita Stroke menggunakan algoritma K-Nearest Neighbor (K-NN) dan Multi-layer Perceptron” yang dapat membantu untuk membantu para masyarakat untuk mengenali apakah orang tersebut mengidap penyakit stroke atau tidak.
4. Terealisasinya “Prediksi Kemungkinan Sesorang Dapat Menderita Stroke menggunakan algoritma K-Nearest Neighbor (K-NN) dan Multi-layer Perceptron” sebagai bentuk kontribusi terhadap *Sustainable Development Goals* (SDGs) terkait kesehatan.

# BAB IV

**PENUTUP**

## Kesimpulan

Dari hasil penelitian yang telat dilakukan dapat disimpulkan, bahwa dalam membuat model machine learning tentang Prediksi Kemungkinan Sesorang Dapat Menderita Stroke bisa menggunakan algoritma K-Nearest Neighbor (K-NN) dan Multi-layer Perceptron untuk pembuatan modelnya. Yang dimana tujuan dari proyek ini yaitu untuk mengklasifikasikan orang-orang yang beresiko terkena penyakit stroke atau tidak. Dari skor yang telah di explorasi dari kedua model tersebut menandakan bahwa hasil prediksi dari model yang dirancang dengan algoritma K-NN memiliki hasil yang lebih baik dari pada model yang dirancang dengan algoritma Multi-layer Perceptron. Dengan adanya model ini diharapkan semua orang dapat mampu menjaga kesehatan agar tidak beresiko terkena penyakit stroke dengan melakukan gaya hidup sehat dan selalu menjaga kesehatan tubuh. Dan juga sebagai sebagai bentuk kontribusi terhadap *Sustainable Development Goals* (SDGs) terkait kesehatan.

## Saran

Untuk MSIB :

* Pengelolaan websitenya harus diperbaiki lagi sebelum melaksanakan pembukaan batch selanjutnya.
* Memberikan informasi yang real dan tidak ada yang membuat kejadian simpang siur antara mahasiswa yang mengikuti kegiatan MSIB.

Untuk Orbit

* Mempersingkat waktu penyampaian materi.
* Memperbanyak waktu untuk melakukan praktik coding.

# REFERENSI

[1] Truelsen, T. dan Begg, S. (2006), The Global Burden Of Cerebrovascular Disease, World Health Organization.

[2] Cardiovascular diseases (CVDs).(2021, Juni 11). who.int. Medium. <https://www.who.int/en/news-room/fact-sheets/detail/cardiovascular-diseases-(cvds)> [Diakses 15 januari 2022].

[3] Infodatin Stroke.(2019, November 29). pusdatin.kemkes.go.id. Medium. <https://pusdatin.kemkes.go.id/article/view/20031000003/infodatin-stroke.html> [Diakses 15 Januari 2022].

[4] Harrison, O. (2019, July 14). Machine Learning Basics with the K-Nearest Neighbors Algorithm. Medium. <https://towardsdatascience.com/machine-learning-basics-with-the-k-nearest-neighbors-algorithm-6a6e71d01761> [Diakses 15 Januari].

[5] Lambrou, A dan Harris, P. (2010). Assessment of Stroke Risk Based on Morphological Ultrasound Image Analysis with Conformal Prediction. IFIP International Federation for Information Processing. No. 146–153.

[6] Akter, S. dan Amin, A. (2021). Stroke prediction analysis using machine learning classifiers and feature technique. International Journal of Electronics and Communications System. Volume 1, Issue 2, 17-22

# LAMPIRAN

## TOR

'Orbit AI Gen Y' adalah program pelatihan Artificial Intelligence holistik yang hybrid dan inklusif untuk siswa dari semua sektor dan minat yang beragam. Program ini menggunakan pendekatan Deep Learning & berbasis proyek.

Orbit AI Gen Y berdurasi 5 bulan dan mencakup mulai dari level basic pengenalan AI, hingga advance level yang mencakup pemrograman Python dan AI Project Cycle (Problem Scoping, Data Acquisition, Data Exploration, Modelling, Evaluation & Deployment). Selain itu program ini juga dilengkapi dengan kursus Life Skills yang penting dan diperlukan untuk kesuksesan perusahaan atau kewirausahaan. Kami memperkuat program ini dengan Sesi Experiential Expert Mentoring dalam Program Arjuna kami. Pembelajar dalam program ini akan bergandeng tangan setiap langkah perjalanannya oleh 16 master mentor kami.

Orbit Future Academy adalah mitra resmi program global Intel AI 4 Youth di 30 negara dengan tujuan untuk menciptakan 30 juta personel yang terlatih dengan AI pada tahun 2030. Oleh karena itu, sertifikat setelah menyelesaikan program ini akan diakui oleh 30.000 organisasi di 30 negara.

## LOG ACTIVITY

Berikut ini adalah log activity dari kegiatan Studi Independen Bersertifikat Kampus Merdeka yang telah dilaksanakan.

|  |  |  |
| --- | --- | --- |
| Minggu/Tgl | Kegiatan | Hasil |
| 01/23-08-2021 | Orientasi Program AI for Gen Y | Gambaran proses belajar di AI4GenY |
| 01/24-08-2021 | Coaching : Pengenalan AI | Pendalaman materi |
| 01/25-08-2021 | Coaching : Pengenalan Domain AI | Pendalaman materi |
| 01/26-08-2021 | Coaching : *System Thinking & Flowchart* | Pendalaman materi |
| 01/27-08-2021 | Bimbingan Teknis dan *Review* Materi | Evaluasi pembelajaran |
| 02/30-08-2021 | Coaching : *Data Visualization* | Pendalaman materi |
| 02/31-08-2021 | Coaching : *AI Ethics and Biases* | Pendalaman materi |
| 02/01-09-2021 | Coaching : Vektor, Skalar & Matriks | Pendalaman materi |
| 02/02-09-2021 | Coaching : *Data is The New Oil* | Pendalaman materi |
| 02/03-09-2021 | Coaching : *Data Presentation* | Pendalaman materi |
| 03/06-09-2021 | Coaching : *AI Project Cycle* | Pendalaman materi |
| 03/07-09-2021 | Coaching : *Data Acquisition* | Pendalaman materi |
| 03/08-09-2021 | Coaching : *Data Exploration* | Pendalaman materi |
| 03/09-09-2021 | Coaching : *AI Modelling* | Pendalaman materi |
| 03/10-09-2021 | Coaching : *AI Evaluation* | Pendalaman materi |
| 04/13-09-2021 | Coaching : *IDE for AI Development* | Pendalaman materi |
| 04/14-09-2021 | Coaching : *Python Data Structures* | Pendalaman materi |
| 04/15-09-2021 | Coaching : *Python OOP* | Pendalaman materi |
| 04/16-09-2021 | Coaching : *Python Libraries for AI Dev* | Pendalaman materi |
| 04/17-09-2021 | Bimbingan Teknis dan *Review* Materi | Evaluasi pembelajaran |
| 05/20-09-2021 | Coaching : *Algorithm & Big O* | Pendalaman materi |
| 05/21-09-2021 | Coaching : Teknik *Debugging* | Pendalaman materi |
| 05/22-09-2021 | Izin | - |
| 05/23-09-2021 | Coaching : *Python OOP Paradigm* | Pendalaman materi |
| 05/24-09-2021 | Bimbingan Teknis dan *Review* Materi | Evaluasi pembelajaran |
| 06/27-09-2021 | Coaching : *KNN Classification* | Implementasi Algoritma KNN |
| 06/28-09-2021 | Coaching : *K-Means Clustering* | Implementasi Algoritma K-Means |
| 06/29-09-2021 | Coaching : *Deep Learning Basics* | Pendalaman materi |
| 06/30-09-2021 | Coaching : *Deep Learning Basics* | Pendalaman materi |
| 06/01-10-2021 | Coaching : *Convolutional Neural Network* | Implementasi Algoritma CNN |
| 07/04-10-2021 | Coaching : *Descriptive Analytics* | Pendalaman materi |
| 07/05-10-2021 | Coaching : CRISP-DM | Pendalaman materi |
| 07/06-10-2021 | Coaching : *Data Preprocessing* | Pendalaman materi |
| 07/07-10-2021 | Coaching : *Data Preprocessing Hands-on* | Implementasi *Data Preprocessing* |
| 07/08-10-2021 | Bimbingan Teknis dan *Review* Materi | Evaluasi pembelajaran |
| 08/11-10-2021 | Coaching : Uji Hipotesis | Pendalaman materi |
| 08/12-10-2021 | Coaching : *Supervised Learning* | Pendalaman materi |
| 08/13-10-2021 | Izin | - |
| 08/14-10-2021 | Coaching : Tableau Data Visualization | Implementasi visualisasi data |
| 08/15-10-2021 | Bimbingan Teknis dan *Review* Materi | Evaluasi pembelajaran |
| 09/18-10-2021 | Coaching : *Natural Language Processing* | Pendalaman materi |
| 09/19-10-2021 | Coaching : *Tools for NLP Development* | Pendalaman materi |
| 09/20-10-2021 | Libur Nasional : Maulid Nabi Muhammad | - |
| 09/21-10-2021 | Coaching : *Scrapping & Cosine Similarity* | Implementasi Scrapping & Cos. Sim. |
| 09/22-10-2021 | Coaching : *Stemming & Lemmatization* | Implementasi Stemming & Lemma |
| 10/25-10-2021 | Coaching : *Sentiment Analysis NLP* | Implementasi *Sentiment Analysis* |
| 10/26-10-2021 | Coaching : *Chatbot with NLP* | Implementasi *Chatbot* |
| 10/27-10-2021 | Coaching : *Text Summarization NLP* | Implementasi *Text Summarization* |
| 10/28-10-2021 | Coaching : *TTS & STT NLP* | Implementasi *TTS & STT* |
| 10/29-10-2021 | Bimbingan Teknis dan *Review* Materi | Evaluasi pembelajaran |
| 11/01-11-2021 | Coaching : *Computer Vision* | Pendalaman materi |
| 11/02-11-2021 | Izin | - |
| 11/03-11-2021 | Izin | - |
| 11/04-11-2021 | Izin | - |
| 11/05-11-2021 | Izin | - |
| 12/08-11-2021 | Coaching : *Image Classification with CV* | Implementasi *Image Classification* |
| 12/09-11-2021 | Coaching : *Image Segmentation with CV* | Implementasi *Image Segmentation* |
| 12/10-11-2021 | Coaching : *Object Detection with CV* | Pendalaman materi |
| 12/11-11-2021 | Coaching : *YOLO v5 Hands On* | Implementasi *Object Detection* |
| 12/12-11-2021 | Coaching : *DGM & GAN Algorithm* | Implementasi DGM & GAN |
| 13/15-11-2021 | Melakukan tahap Problem Scoping, yakni menentukan permasalahan yang ingin dipecahkan dalam penelitian ini dengan melihat isu dan topik-topik populer yang sedang berkembang di lingkungan masyarakat. Selain melakukan tahap problem scoping, saya mengikuti kelas coaching yang membahas tentang beberapa corporate skills yang bermanfaat untuk pengembangan diri. | Literasi terkait isu-isu populer |
| 13/16-11-2021 | Literasi terkait isu-isu populer |
| 13/17-11-2021 | Literasi terkait isu-isu populer |
| 13/18-11-2021 | Literasi terkait isu-isu populer |
| 13/19-11-2021 | Literasi terkait isu-isu populer |
| 14/22-11-2021 | Literasi terkait isu-isu populer |
| 14/23-11-2021 | Literasi terkait isu-isu populer |
| 14/24-11-2021 | Literasi terkait isu-isu populer |
| 14/25-11-2021 | Penentuan pokok permasalahan |
| 14/26-11-2021 | Penentuan pokok permasalahan |
| 15/29-11-2021 | Melakukan eksplorasi pada paper-paper penelitian terkait. | Terkumpulnya paper terkait |
| 15/30-11-2021 | Terpilihnya algoritma yang sesuai |
| 15/01-12-2021 | Melakukan tahap Data Acquisition, yakni pengumpulan data yang akan digunakan pada model AI. Dataset didapatkan dari situs kaggle. | Dataset dari kaggle didapatkan |
| 15/02-12-2021 | Dataset dari kaggle didapatkan |
| 15/03-12-2021 | Dataset dari kaggle didapatkan |
| 16/06-12-2021 | Dataset dari kaggle didapatkan |
| 16/07-12-2021 | Pemilihan dataset terbaik |
| 16/08-12-2021 | Melakukan tahap Data Exploration, yakni menganalisis data yang telah dikumpulkan dari berbagai sumber. Analisis dilakukan untuk mengetahui struktur dan karakteristik dari dataset yang telah didapatkan pada tahap data acquisition. Hasil analisis nantinya akan digunakan sebagai modal utama dalam melakukan penanganan di tahap *data cleaning*, serta tahap *data preprocessing* lainnya. | Analisis struktur dataset |
| 16/09-12-2021 | Analisis struktur dataset |
| 16/10-12-2021 | Analisis struktur dataset |
| 17/13-12-2021 | Analisis struktur dataset |
| 17/14-12-2021 | Analisis struktur dataset |
| 17/15-12-2021 | Analisis struktur dataset |
| 17/16-12-2021 | Analisis struktur dataset |
| 17/17-12-2021 | Analisis struktur dataset |
| 18/20-12-2021 | Analisis struktur dataset |
| 18/21-12-2021 | Analisis struktur dataset |
| 18/22-12-2021 | Data cleaning : Missing Data Handling | Pembersihan *Missing Data* |
| 18/23-12-2021 | Data cleaning : Inconsistent Data Handling | Pembersihan *Inconsistent Data* |
| 18/24-12-2021 | Data cleaning : Outliers Data Handling | Pembersihan *Outliers Data* |
| 19/27-12-2021 | Data cleaning : Redundant Data Handling | Pembersihan *Redundant Data* |
| 19/28-12-2021 | Membagi dataset *training* dan *tersting* | Pembagian data *train* dan *test* |
| 19/29-12-2021 | Menyusun arsitektur K-NN dan MLP | Penyusunan arsitektur *K-NN & MLP* |
| 19/30-12-2021 | Mentransformasikan dataset | Transformasi data |
| 19/31-12-2021 | Melakukan finishing transformasi dataset | Finalisasi transformasi data |
| 20/03-01-2022 | Melakukan tahap Modelling, yakni merancang arsitektur dari model AI yang akan digunakan untuk memprediksi seseorang terkena stroke atau tidak selama satu minggu setelah waktu terakhir di dataset. | Terbentuknya struktur model |
| 20/04-01-2022 | Dilakukannya proses training |
| 20/05-01-2022 | Dilakukannya proses training |
| 20/06-01-2022 | Dilakukannya proses training |
| 20/07-01-2022 | Dilakukannya proses testing |
| 21/10-01-2022 | Melakukan tahap Evaluation terhadap performa model, serta mendapatkan data hasil prediksi. | *Performance metrics* didapatkan |
| 21/11-01-2022 | *Performance metrics* didapatkan |
| 21/12-01-2022 | *Performance metrics* didapatkan |
| 21/13-01-2022 | *Performance metrics* didapatkan |
| 21/14-01-2022 | Analisis performa model |
| 22/17-01-2022 | Analisis performa model |
| 22/18-01-2022 | Data hasil prediksi |
| 22/19-01-2022 | Penyusunan Laporan Akhir Kegiatan MBKM – Studi Independen Batch 1 | Rancangan laporan |
| 22/20-01-2022 | Rancangan laporan |
| 22/21-01-2022 | Rancangan laporan |
| 22/22-01-2022 | Rancangan laporan |
| 22/23-01-2022 | Rancangan laporan |
| 22/24-01-2022 | Rancangan laporan |
| 22/25-01-2022 | Rancangan laporan |
| 22/26-01-2022 | Rancangan laporan |
| 22/27-01-2022 | Finalisasi rancangan laporan |
| 22/28-01-2022 | Finalisasi rancangan laporan |
| 22/29-01-2022 | Finalisasi rancangan laporan |

## Dokumen Teknik

### AI Project Cycle

* + 1. Problem Scoping

Stroke adalah kondisi yang terjadi ketika pasokan darah ke otak terganggu atau berkurang akibat penyumbatan (stroke iskemik) atau pecahnya pembuluh darah (stroke hemoragik). Tanpa darah, otak tidak akan mendapatkan asupan oksigen dan nutrisi, sehingga sel-sel pada sebagian area otak akan mati. Kondisi ini menyebabkan bagian tubuh yang dikendalikan oleh area otak yang rusak tidak dapat berfungsi dengan baik. Faktor yang menyebabkan terjadinya stroke antara lain adalah faktor kesehatan, faktor gaya hidup, faktor usia bahkan faktor keturunan.

Diperkirakan 17,9 juta orang meninggal karena CVD (Penyakit kardiovaskular) pada 2019, mewakili 32% dari semua kematian global. Dari kematian tersebut, 85% disebabkan oleh serangan jantung dan stroke menurut laporan dari WHO (The World Health Organization)[2]. Berdasarkan laporan dari Pusat Data dan Informasi Kementrian Kesehatan RI, secara nasional prevalensi stroke di Indonesia tahun 2018 berdasarkan diagnosis dokter pada penduduk berumur diatas 15 tahun sebesar 10,9%, atau diperkirakan sebanyak 2.120.362 orang[3 ]. Maka dari itu diperlukan kesadaran bagi semua orang untuk menghindari resiko-resiko yang dapat menyebabkan penyakit stroke. Untuk menyadarkan semua orang akan bahaya stroke perlu adanya keterlibatan teknologi modern sebagai pencegahan penderita mengalami stroke yang tiba-tiba. Salah satunya menggunakan model Machine Learning untuk menginformasi lebih awal tanda-tanda seseorang akan terkena stroke atau tidak.

* + 1. Data Acquisition

Sebagai bahan untuk melakukan analisis menggunakan AI, tentu kita memerlukan dataset stroke yang telah ada guna sebagai informasi awal machine. Data yang kami gunakan berasal dari situs website, yakni Kaggel. Melalui situs web ini, kami mendapatkan data stroke yang cukup lengkap dengan banyak data sebanyak 44rb lebih data stroke.

* + 1. Data Exploration

Setelah mengumpulkan data dari situs kaggle, tentu perlu dilakukan analisis sebelum akhirnya data masuk ke tahap modelling. Analisis dilakukan melalui google colab untuk memudahkan dalam tahap explorasi data. Setelah dilakukan analisis, ternyata dataset yang didapatkan masih mengandung berbagai noise, seperti seperti *missing data*, *outliers*, dan lain sebagai. Pembersihan terhadap data noise dilakukan dengan melakukan pembersih dengan code python dan library pendukung yang tersedia pada python. Setelah dataset bersih, data siap digunakan pada tahap modelling.

* + 1. Modelling

Dalam mengklasifikasi Kemungkinan Sesorang Dapat Menderita Stroke mengggunakan data yang ada, peneliti menggunakan K-Nearest Neighbor dan Multi-layer Perceptron. Pemilihan algoritma ini didasarkan pada dataset yang dimiliki peneliti memiliki data berkategori dan data numerik sehingga cocok menggunakan 2 algoritma tersebut. Implementasi algoritma tersebut dilakukan melalui Google Colab dengan library dari scikit-learn. Model yang dibuat berupa klasifikasi dengan 2 algoritma yang digunakan peneliti.

* + 1. Evaluation

Pengujian performanya dilakukan dengan menggunakan matrikx akurasi, f1-score, recall dan precision. Setelah dilakukan pengujian, model yang dirancang menggunakan K-NN menghasikan skor akurasi 89%, precision 90%, recall 89%, dan f1-score 89%. Sedangkan model yang dirancang menggunakan Multi-layer Perceptron menghasilkan skor akurasi 88%, precision 88%, recall 88%, dan f1-score 88%. Skor yang ditunjukkan menandakan bahwa hasil prediksi dari model yang dirancang dengan algoritma K-NN memiliki hasil yang lebih baik dari pada model yang dirancang dengan algoritma Multi-layer Perceptron.

### Profil Peneliti

* 1. **Muhamad Rendi**
* Asal Kampus : UNIVERSITAS TEKNOLOGI YOGYAKARTA
* Jurusan : Informatika
* Email : [muh.rendi2001@gmail.com](mailto:muh.rendi2001@gmail.com)
* No.Hp : +62 89670128440