

# Pemodelan Klasifikasi Penyakit Alzheimer

Tri Indah Wahyuningsi



# Tentang Saya

Saya Tri Indah Wahyuningsi, lulusan Teknik Informatika dari Universitas Hasanuddin. Saat ini saya memiliki ketertarikan yang besar di bidang Data Science dan terus mengembangkan kemampuan melalui pembelajaran mandiri, pelatihan/bootcamp, serta pengerjaan berbagai proyek skala kecil. Fokus saya adalah membangun kompetensi dalam analisis data dan pengolahan informasi untuk menghasilkan solusi yang efektif dan relevan dalam proses pengambilan keputusan.



# Pengalaman Kerja

**PT SAMATA GREENHOUSE**

*Dokumentasi Teknis (Agustus 2023 – Desember 2023)*

Berkontribusi dalam proyek pengembangan sistem pertanian hidroponik cerdas melalui program MBKM Kedaireka 2023 sebagai dokumentasi teknis yang berperan sebagai analis kebutuhan sistem, menyusun kebutuhan administrasi, membantu tim konfigurasi sensor dalam pemasangan dan kalibrasi beberapa sensor yaitu (pH, TDS, DHT, level air dan suhu air).

# Pendidikan



**UNIVERSITAS HASANUDDIN  
SARJANA TEKNIK INFORMATIKA (2020-2024)**

Merancang sistem kendali otomatis kumbung jamur pada budidaya jamur tiram menggunakan fuzzy logic untuk mengatur kondisi lingkungan dalam kumbung agar tetap ideal.

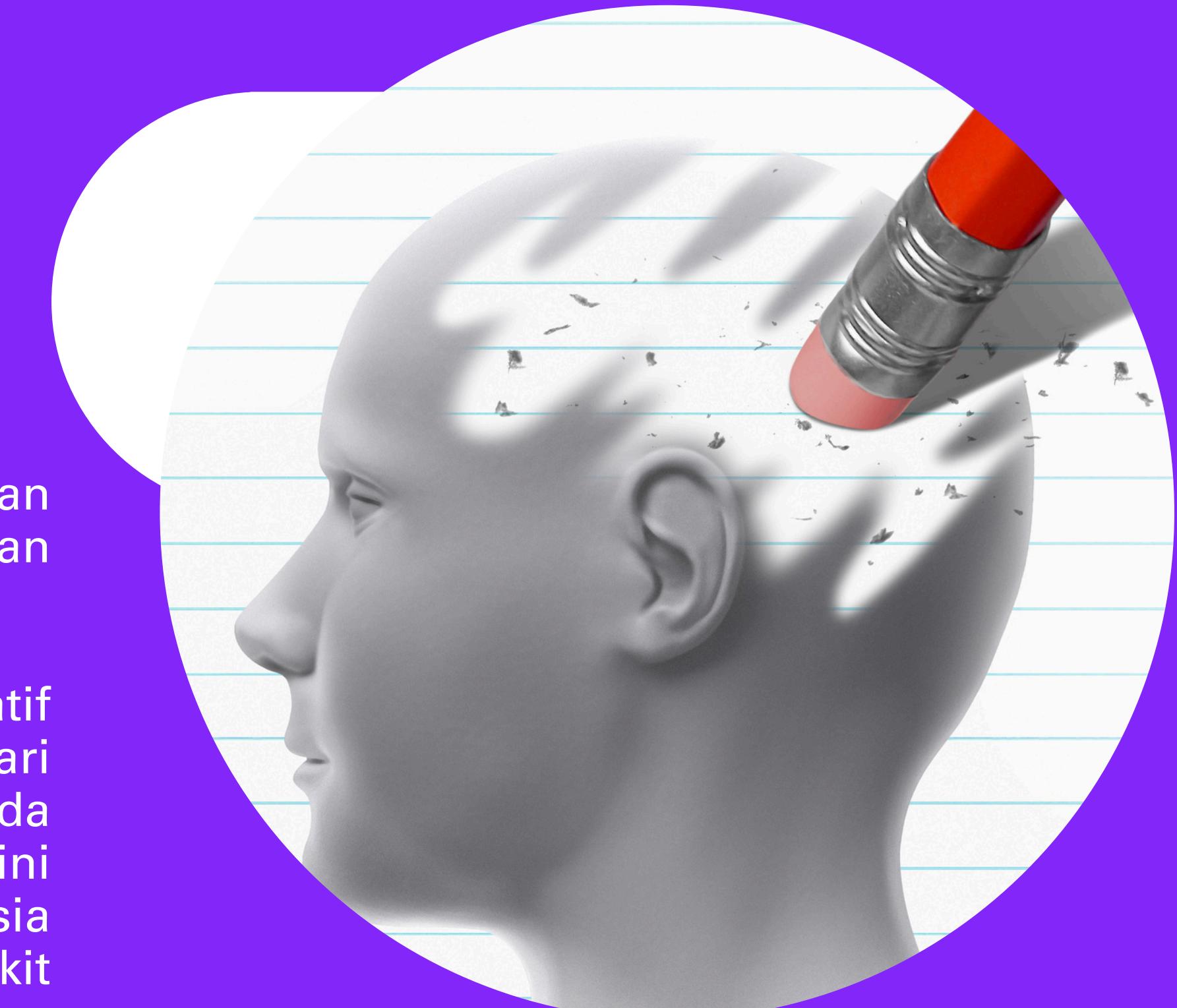


**BOOTCAMP DATA SCIENCE & DATA  
ANALYST (MEI 2025 - SEKARANG)**

Mempelajari mengenai konsep fundamental data science, manajemen database, pembuatan dashboard, hingga pemodelan machine learning.

# Gangguan Otak: Demensia, Alzheimer

- Demensia adalah kondisi hilang ingatan yang timbul bersama gejala gangguan perilaku dan psikologis pada seseorang.
- Alzheimer adalah penyakit degeneratif otak dan penyebab paling umum dari demensia. Penyakit Alzheimer terjadi pada setiap dekade dewasa, tapi penyakit ini merupakan penyebab utama demensia pada lanjut usia (Bahan Ajar 1 Penyakit Alzheimer, 2013).



SUMBER : MAJALAH KESEHATAN  
INDONESIA, 2021



## Tanda-tanda Penyakit Alzheimer

- Penurunan memori, bahasa, pemecahan masalah, dan keterampilan kognitif lainnya yang memengaruhi seseorang dalam melakukan aktivitas sehari-hari. Penurunan ini terjadi karena sel-sel saraf di bagian otak yang terlibat dalam fungsi kognitif telah rusak dan tidak lagi berfungsi normal (Alzheimer's Association, 2017).
- Terlihat adanya bukti **otak menyusut (atrofi)** yang bertambah parah dari waktu ke waktu (dilihat dari hasil *scan* yang dilakukan berulang kali).

# Memahami Data

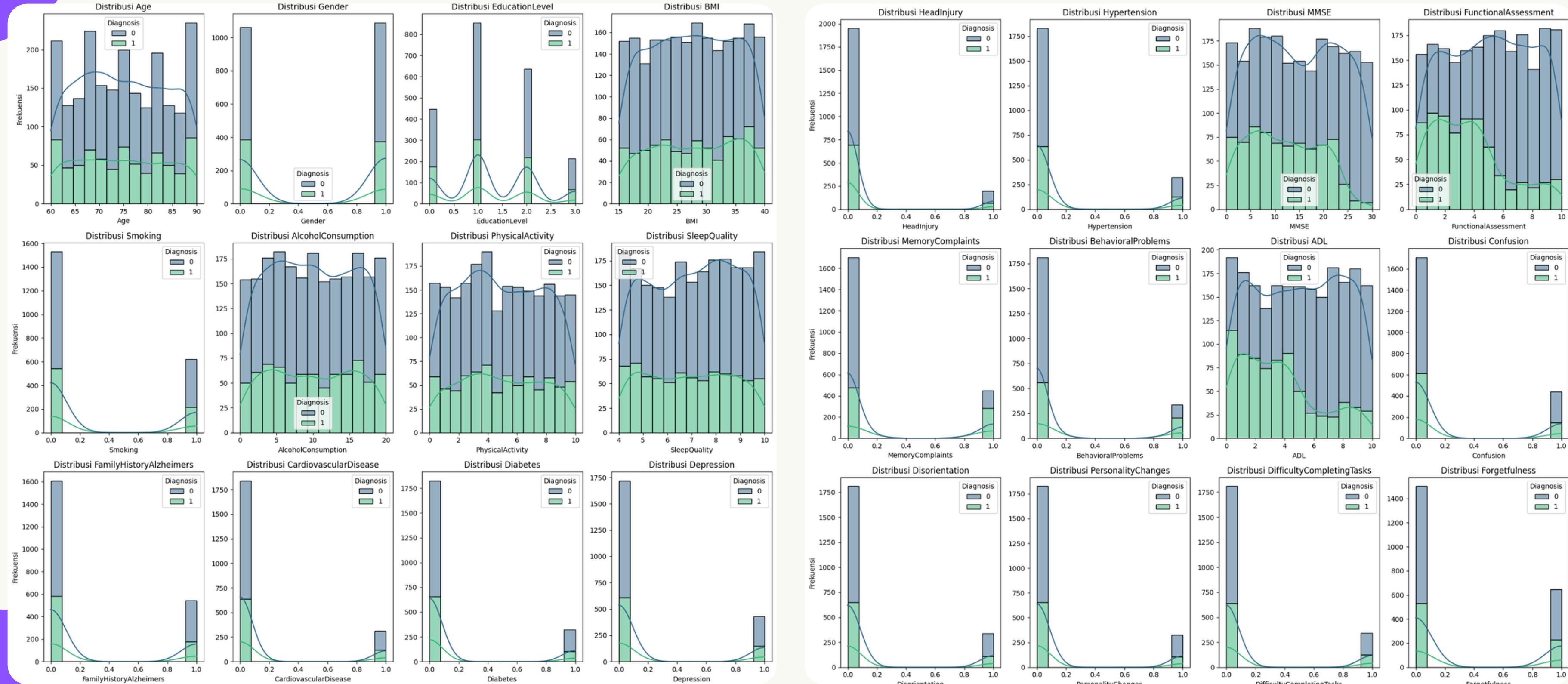
- ✓ **Dataset ini diperoleh melalui platform Kaggle**
- ✓ **Dataset ini memiliki 2.149 pasien dengan variabel ukur sebanyak 35 variabel**
- ✓ **Tidak ada data duplikat dan data yang hilang**
- ✓ **Outlier data masih wajar**
- ✓ **Melakukan seleksi variabel**

# Analisis Objektif

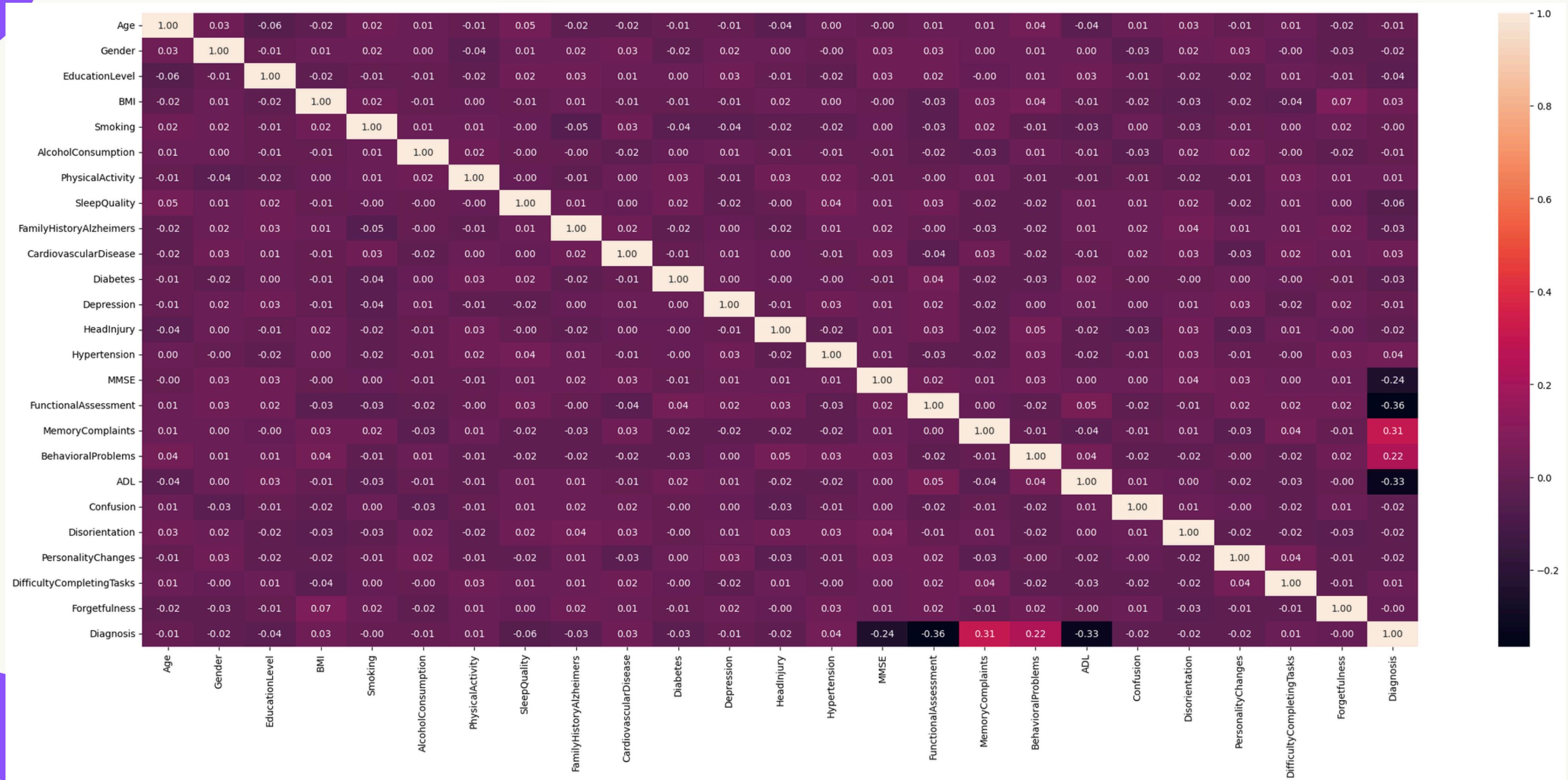
Menganalisis faktor-faktor yang memungkinkan terkena Alzheimer dengan eksplorasi analisis data (EDA) berdasarkan hasil pemeriksaan kognitif, rekam medis, faktor gaya hidup pasien, gejala yang dialami hingga informasi demografi pasien untuk mengidentifikasi kontributor utama dalam memprediksi diagnosis penyakit ini.



# Distribusi Data



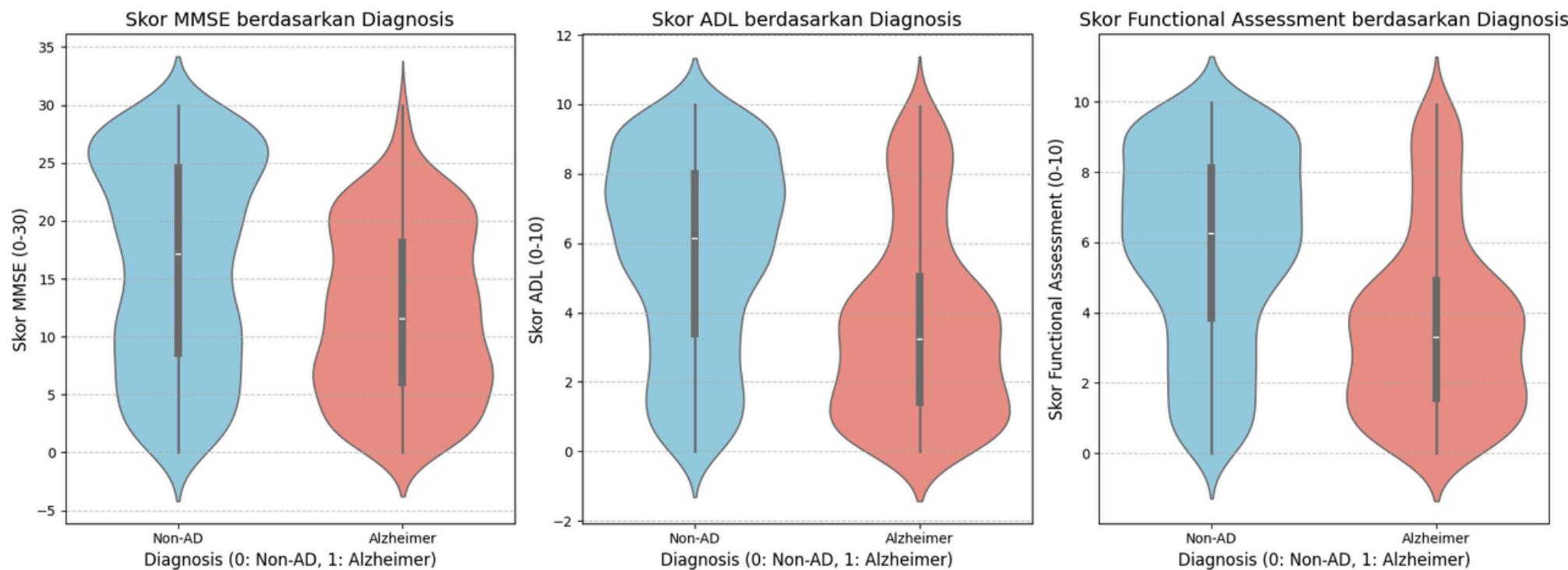
# Korelasi antara Variabel



## Insight : Korelasi antara Variabel

1. MMSE (-0.24) : Menunjukkan korelasi negatif. Artinya, semakin rendah skor Mini-Mental State Examination, semakin tinggi kemungkinan pasien menderita Alzheimer. Secara klinis bahwa Alzheimer menyebabkan penurunan kemampuan kognitif yang terukur melalui skor MMSE.
2. Functional Assessment (-0.36) : Menunjukkan korelasi negatif. Artinya, pasien dengan penilaian fungsional rendah (FunctionalAssessment rendah) memiliki risiko lebih besar terkena Alzheimer.
3. ADL (-0.33) :Menunjukkan korelasi negatif. Artinya, semakin menurun kemampuan pasien dalam menjalankan aktivitas harian, semakin besar kemungkinan mereka memiliki Alzheimer.
4. Memory Complaints (0.31) : Menunjukkan korelasi positif. Artinya, pasien dengan keluhan memori cenderung memiliki diagnosis Alzheimer lebih tinggi. Secara klinis, gejala utama Alzheimer yaitu penurunan daya ingat jangka pendek dan panjang.
5. Behavioral Problems (0.22) : Menunjukkan korelasi positif. Artinya, pasien yang mengalami masalah perilaku memiliki kecenderungan lebih tinggi terhadap diagnosis Alzheimer. Secara klinis, gangguan perilaku seperti mudah marah, cemas, atau perubahan kepribadian sering muncul pada tahap menengah Alzheimer.
6. Age, Gender, BMI, AlcoholConsumption, SleepQuality, CardiovascularDisease, Diabetes, Depression, HeadInjury, Hypertension, dan PhysicalActivity (rentang -0.06 hingga 0.05): Korelasi yang sangat lemah tersebut menunjukkan bahwa faktor-faktor tersebut tidak memiliki hubungan linear langsung dengan diagnosis Alzheimer

# Apakah rata-rata skor MMSE, Functional Assessment, dan ADL pasien Alzheimer secara signifikan lebih rendah daripada Non-AD?



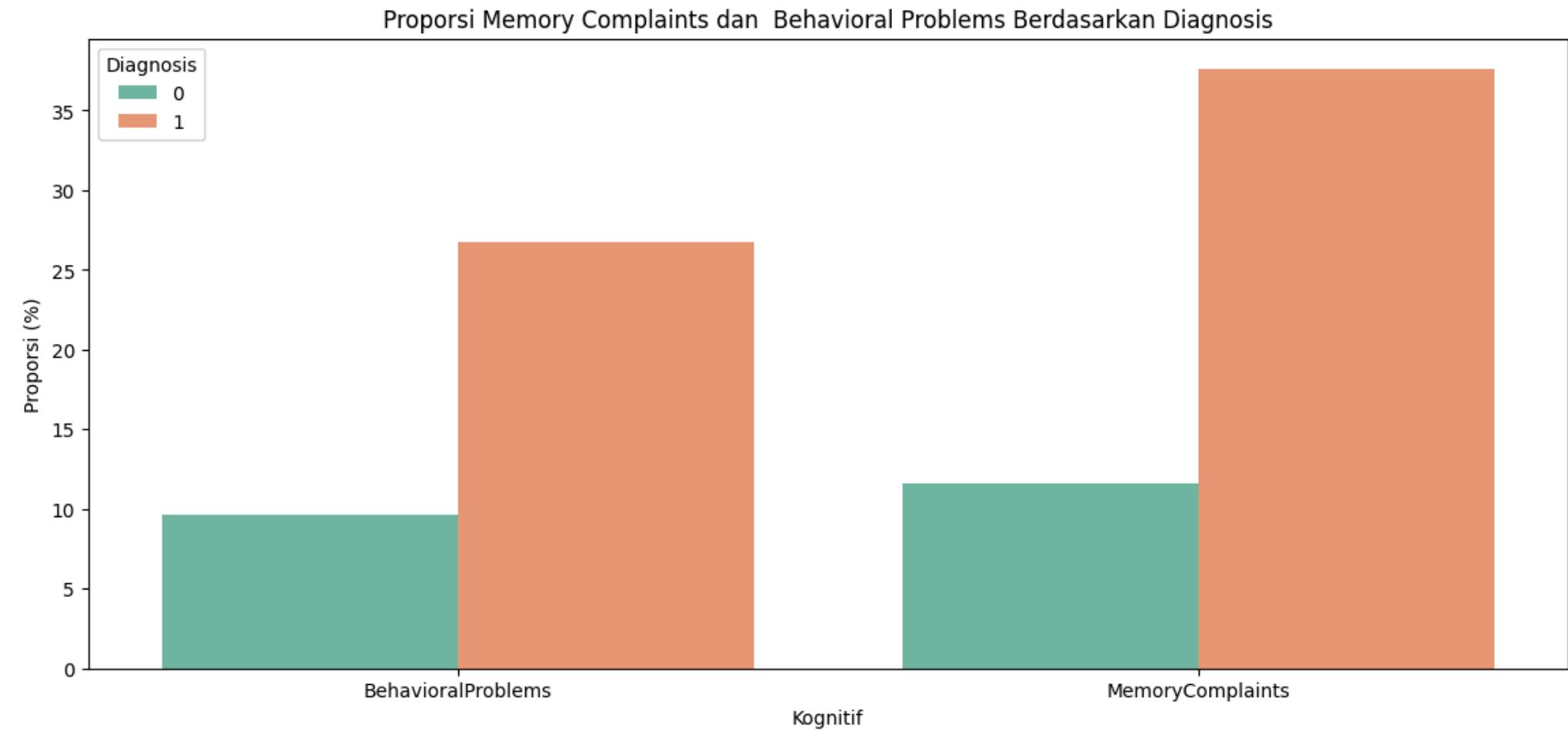
## HASIL UJI T-TEST

Variabel	T-Statistic	P-value	Signifikansi
MMSE	12.025	0.000	Signifikan
ADL	16.546	0.000	Signifikan
Functional Assessment	18.552	0.000	Signifikan

- Rata-rata skor MMSE, Functional Assessment, dan ADL pada pasien Alzheimer secara signifikan lebih rendah dibandingkan pasien Non-AD.
- Hal ini menunjukkan bahwa kemampuan kognitif (MMSE), fungsi aktivitas harian (ADL), functional assessment menurun secara nyata pada pasien yang diagnosis Alzheimer.
- Perbedaan distribusi pada violin plot juga mendukung hasil ini, dapat dilihat bahwa sebaran skor pasien Alzheimer cenderung bergeser ke nilai yang lebih rendah dibanding Non-AD.

# Apakah Behavioral Problems dan Memory Complaints sering terjadi pada pasien Alzheimer?

## HASIL UJI CHI-SQUARE

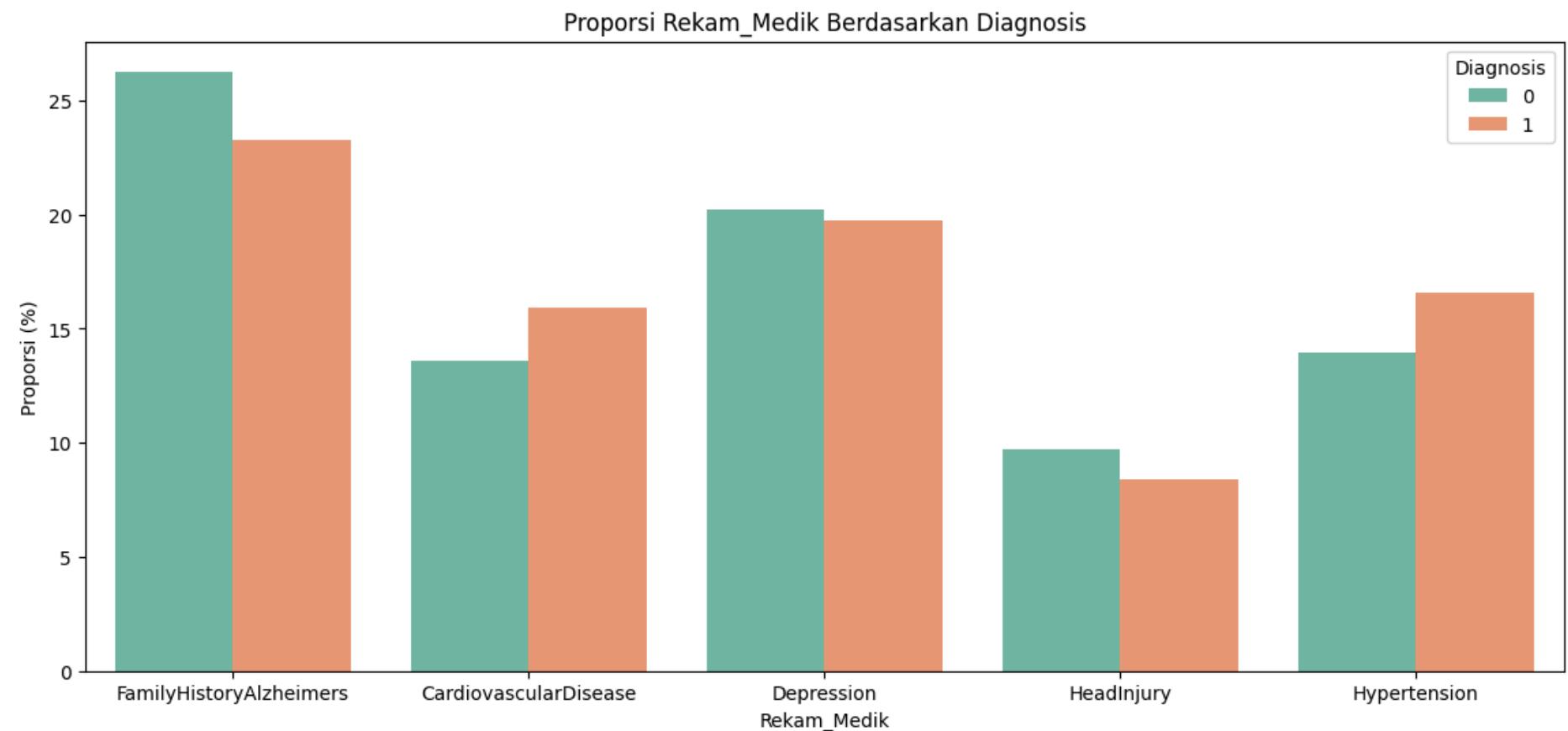


Variabel	Chi2	P-Value	Signifikansi
Behavioral Problems	106.879	0.000	Signifikan
Memory Complaints	200.623	0.000	Signifikan

- Pasien dengan keluhan memory complaints dan behavioral problems memiliki proporsi yang secara signifikan lebih tinggi pada kelompok Alzheimer
- Hasil uji chi-square menunjukkan bahwa kedua variabel tersebut berhubungan signifikan dengan diagnosis Alzheimer ( $p\text{-value} < 0.05$ ), menandakan bahwa gangguan memori dan perubahan perilaku merupakan indikator kognitif dan emosional penting dalam identifikasi Alzheimer.

# Apakah proporsi pasien dengan riwayat medis seperti Family History of Alzheimer's, Cardiovascular Disease, Depression, Hypertension, dan Head Injury lebih tinggi pada kelompok Alzheimer dibandingkan Non-AD?

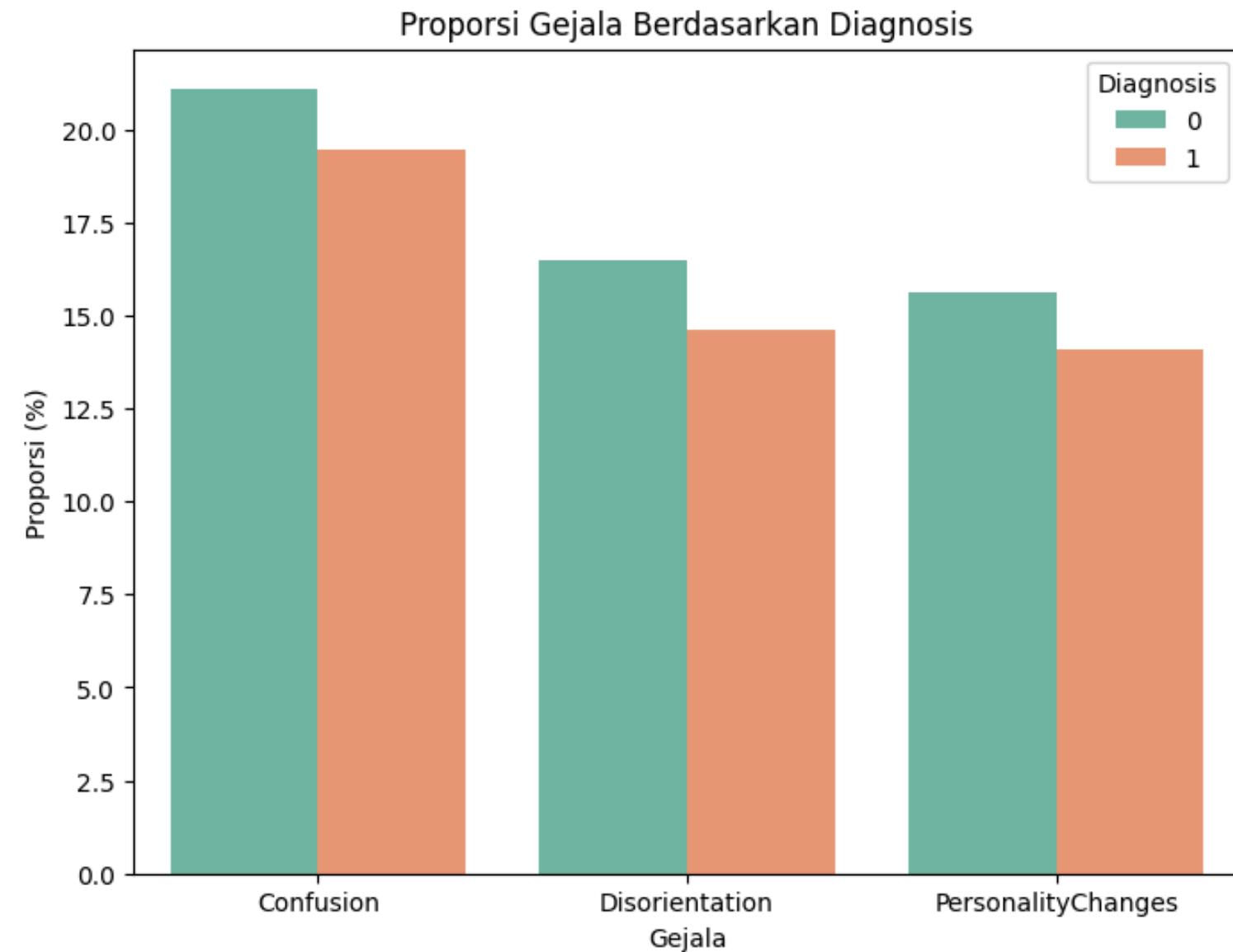
## HASIL UJI CHI-SQUARE



Variabel	Chi2	P-value	Signifikansi
Family History Alzheimers	2.170	0.140	Tidak signifikan
Cardiovascular Disease	1.947	0.162	Tidak Signifikan
Depression	0.470	0.826	Tidak Signifikan
Hypertension	2.442	0.118	Tidak Signifikan
Head Injury	0.836	0.360	Tidak Signifikan

Riwayat medis (Family History, Depression, Head Injury) memiliki proporsi lebih tinggi pada Non-Alzheimer dan proporsi pasien Alzheimer lebih tinggi pada riwayat cardiovascular disease dan hypertension. Namun, berdasarkan hasil uji chi-square menunjukkan tidak ada hubungan signifikan antara riwayat medis dan diagnosis Alzheimer.

# Gejala mana yang sering terjadi pada pasien Alzheimer, dan manakah yang paling jarang terjadi di antara Non-AD?

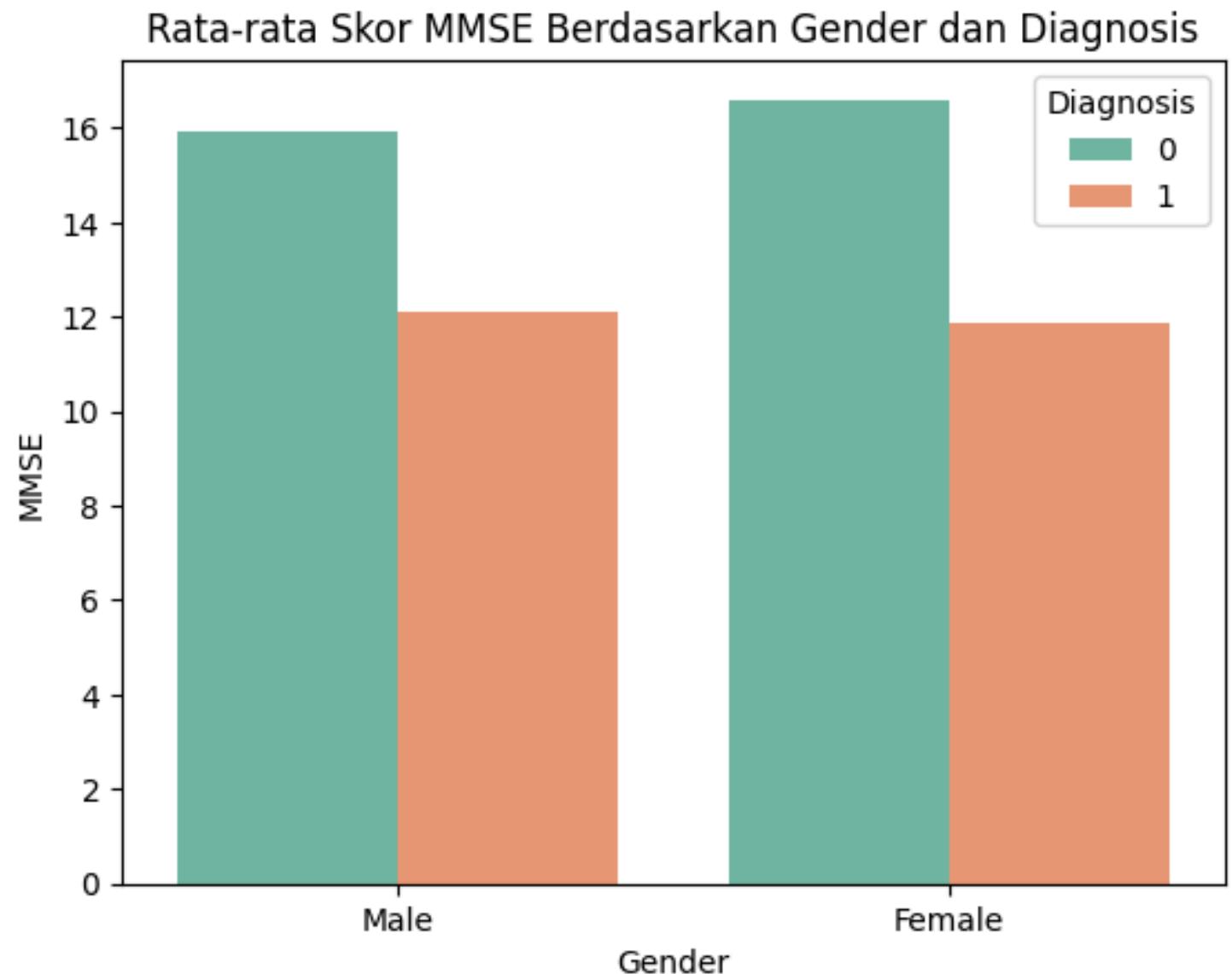


## HASIL UJI CHI-SQUARE

Variabel	Chi2	P-Value	Signifikansi
Confusion	0.694	0.404	Tidak Signifikan
Disorientation	1.168	0.279	Tidak Signifikan
Personality Changes	0.797	0.371	Tidak Signifikan

- Gejala confusion sebagai salah satu faktor yang dapat diidentifikasi pasien tersebut menderita Alzheimer atau tidak, disusul gejala disorientation. Namun, tidak memiliki hubungan secara signifikan terhadap diagnosis Alzheimer

# Bagaimana perbandingan rata-rata MMSE antar Gender untuk pasien Alzheimer dan Non-AD?

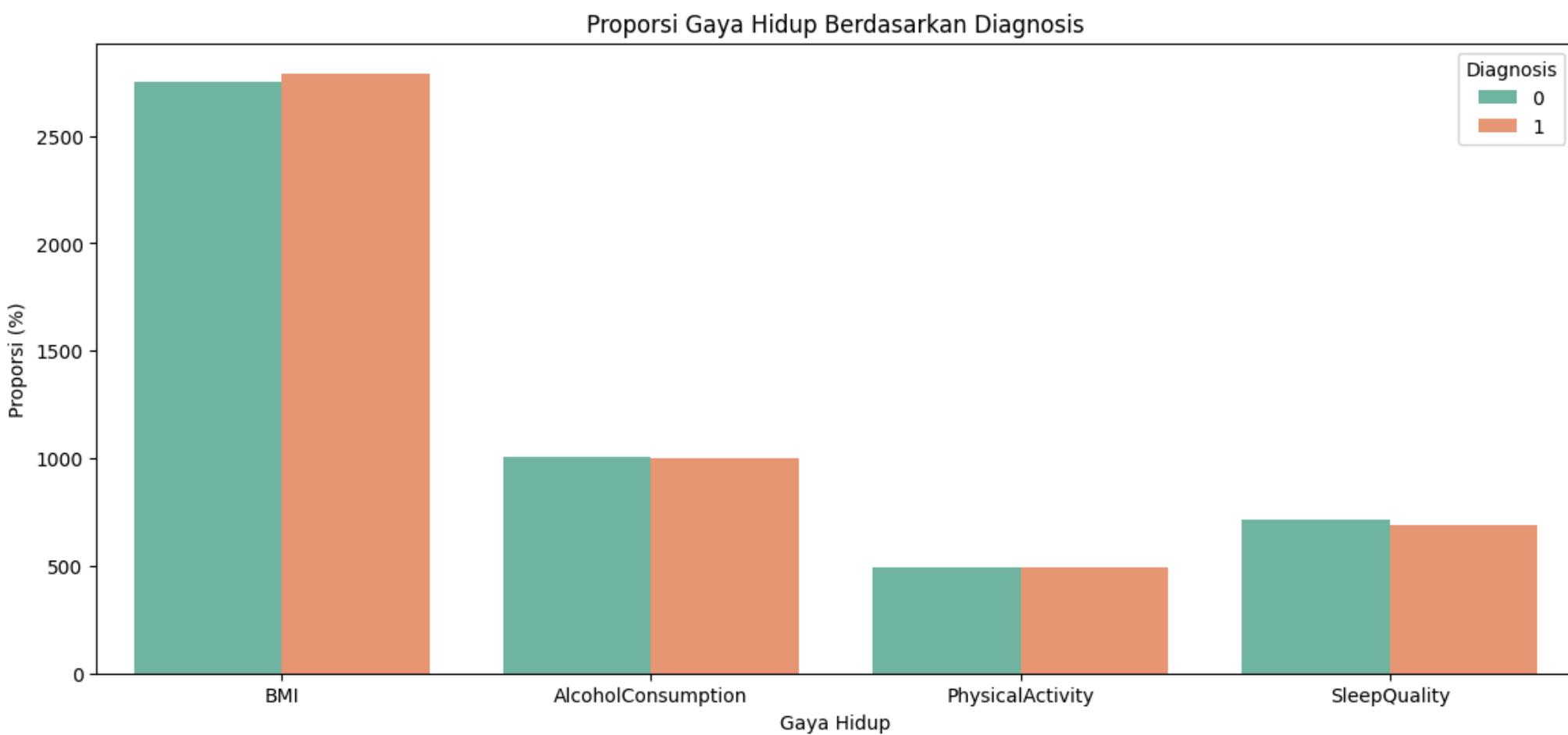


PROPORSI

Gender	Diagnosis	Value(%)
0	0	15.911
	1	12.126
1	0	16.600
	1	11.858

- Pada pasien laki-laki dan perempuan menunjukkan kemampuan kognitif yang menurun pada diagnosis Alzheimer
- Penurunan skor ini cenderung hampir merata baik pada pasien laki-laki dan perempuan. Namun, pada pasien laki laki memiliki skor MMSE lebih tinggi daripada perempuan. Artinya, laki-laki cenderung memiliki fungsi kognitif yang sedikit lebih baik dibandingkan perempuan di dataset ini.

# Bagaimana perbandingan proporsi gaya hidup antara pasien Alzheimer dan Non-AD?



## HASIL UJI CHI-SQUARE

Variabel	Chi2	P-Value	Signifikansi
BMI	2149.0	0.489	Tidak Signifikan
Alcohol Consumption	2149.0	0.489	Tidak Signifikan
Physical Activity	2149.0	0.489	Tidak Signifikan
Sleep Quality	2149.0	0.489	Tidak Signifikan

- Berdasarkan hasil analisis, BMI memiliki proporsi tertinggi diantara gaya hidup lain. Namun, selisih rata-rata BMI pada pasien Alzheimer dan Non-AD sangat kecil. Faktor gaya hidup lainnya menunjukkan proporsi yang cenderung mirip antara pasien Alzheimer dan non-AD
- Hasil uji chi-square menunjukkan bahwa tidak terdapat hubungan secara signifikan pada semua faktor gaya hidup pada diagnosis Alzheimer

# Rekomendasi

- ✓ Melakukan pemeriksaan secara rutin
- ✓ Membuat model prediksi sebagai deteksi dini terhadap diagnosa Alzheimer
- ✓ Edukasi pola hidup sehat dan program aktivitas fisik pada lansia
- ✓ Terapi psikologi dan pendampingan perilaku

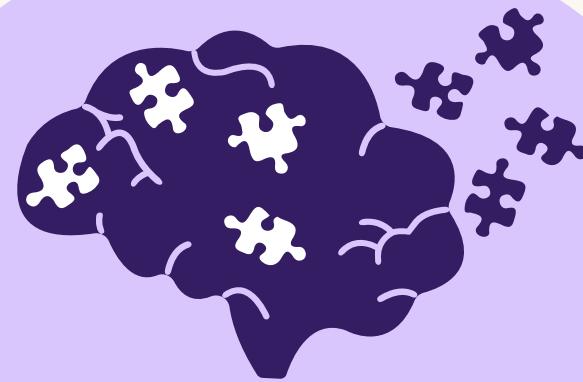


# Kesimpulan

- Diagnosis Alzheimer lebih dipengaruhi oleh faktor kognitif dan perilaku dibandingkan faktor medis atau gaya hidup, sesuai dengan karakteristik klinis penyakit ini.
- Hal ini ditunjukkan oleh hasil analisis yang menemukan bahwa rata-rata skor MMSE, ADL, dan Functional Assessment pada pasien Alzheimer secara signifikan lebih rendah dibandingkan pasien Non-AD.
- Selain itu, proporsi pasien dengan keluhan kognitif (Memory Complaints) dan masalah perilaku (Behavioral Problems) juga secara signifikan lebih tinggi pada kelompok Alzheimer, sehingga kedua gejala ini dapat dianggap sebagai indikator penting dalam deteksi dan klasifikasi penyakit diagnosis Alzheimer.



# Data Preprocessing



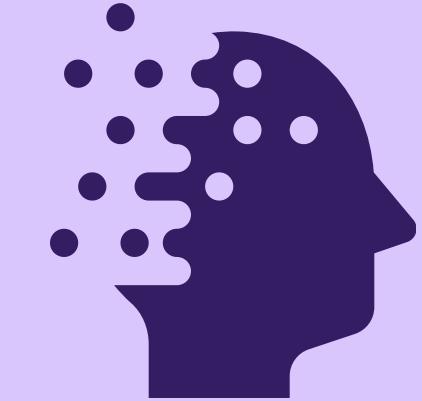
## Split Data

Split data dengan rasio 70 : 30



## Cek Multicollinearity

Cek koefisien tiap variabel  
untuk melihat besaran korelasi  
yang diperoleh



## Scaling

Menormalisasikan data  
menggunakan min max scaler

# Pemodelan Klasifikasi



# Pemodelan Klasifikasi (1)

- Rasio split data 70 : 30
- Menggunakan Min Max Scaler
- Membuang kolom 'Smoking','Forgetfulness', 'Ethnicity', 'DietQuality', 'PatientID', 'SystolicBP', 'DiastolicBP','DoctorInCharge','CholesterolTotal','CholesterolLDL', 'CholesterolHDL', 'CholesterolTriglycerides'

	Model	Precision (train)	Recall (train)	F1-Score (train)	ROC AUC (train)	Precision (test)	Recall (test)	F1-Score (test)	ROC AUC (test)
0	Logistic Regression	0.731667	0.850775	0.786738	0.914682	0.718045	0.782787	0.749020	0.882660
1	Decision Tree	1.000000	1.000000	1.000000	1.000000	0.866667	0.852459	0.859504	0.886329
2	Random Forest	1.000000	1.000000	1.000000	1.000000	0.946341	0.795082	0.864143	0.946016
3	XGBoost	1.000000	1.000000	1.000000	1.000000	0.946667	0.872951	0.908316	0.946037
4	LightGBM	1.000000	1.000000	1.000000	1.000000	0.951327	0.881148	0.914894	0.945444
5	KNN	0.838164	0.672481	0.746237	0.909803	0.685535	0.446721	0.540943	0.746714
6	SVM	0.843860	0.932171	0.885820	0.961958	0.738281	0.774590	0.756000	0.885021
7	MLP	0.838926	0.242248	0.375940	0.803261	0.780822	0.233607	0.359621	0.802747

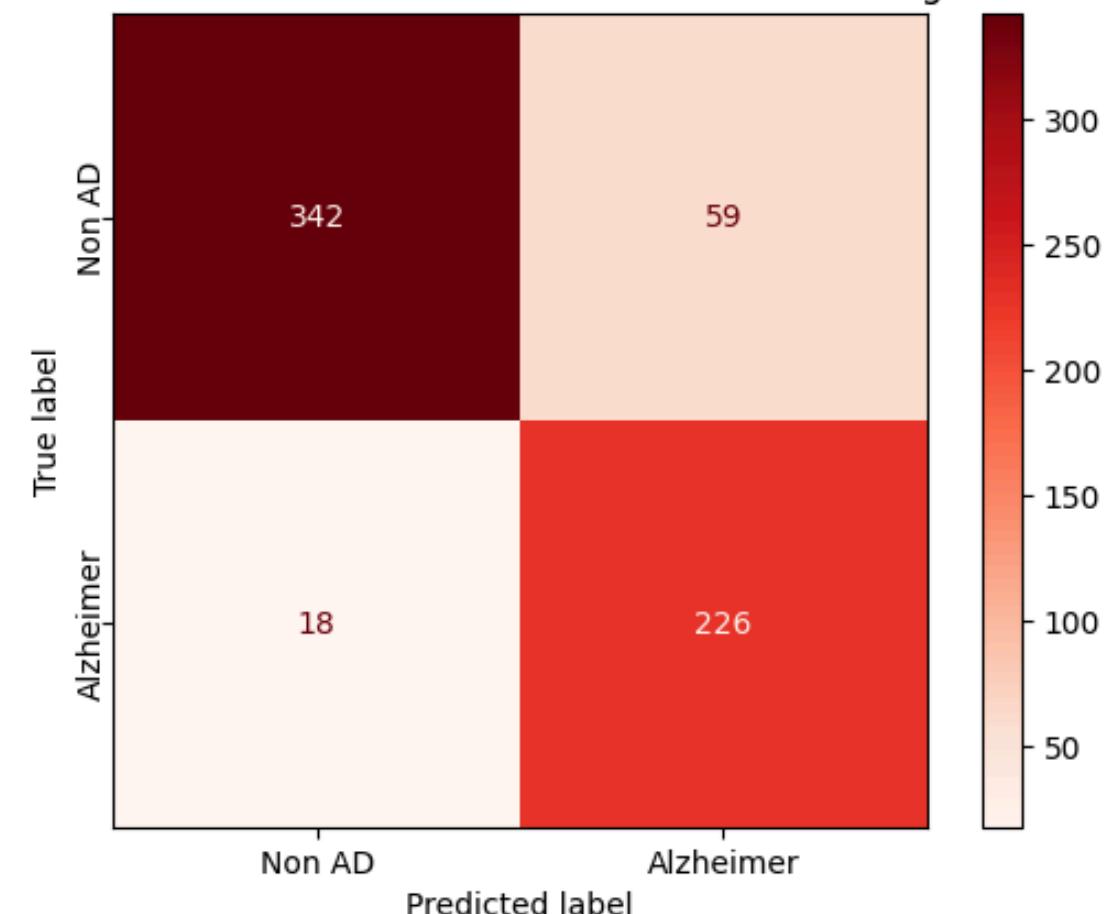
# Pemodelan Klasifikasi (1)

- Hyperparameter Tuning
- Hyperparameter Tuning with SMOTEENN

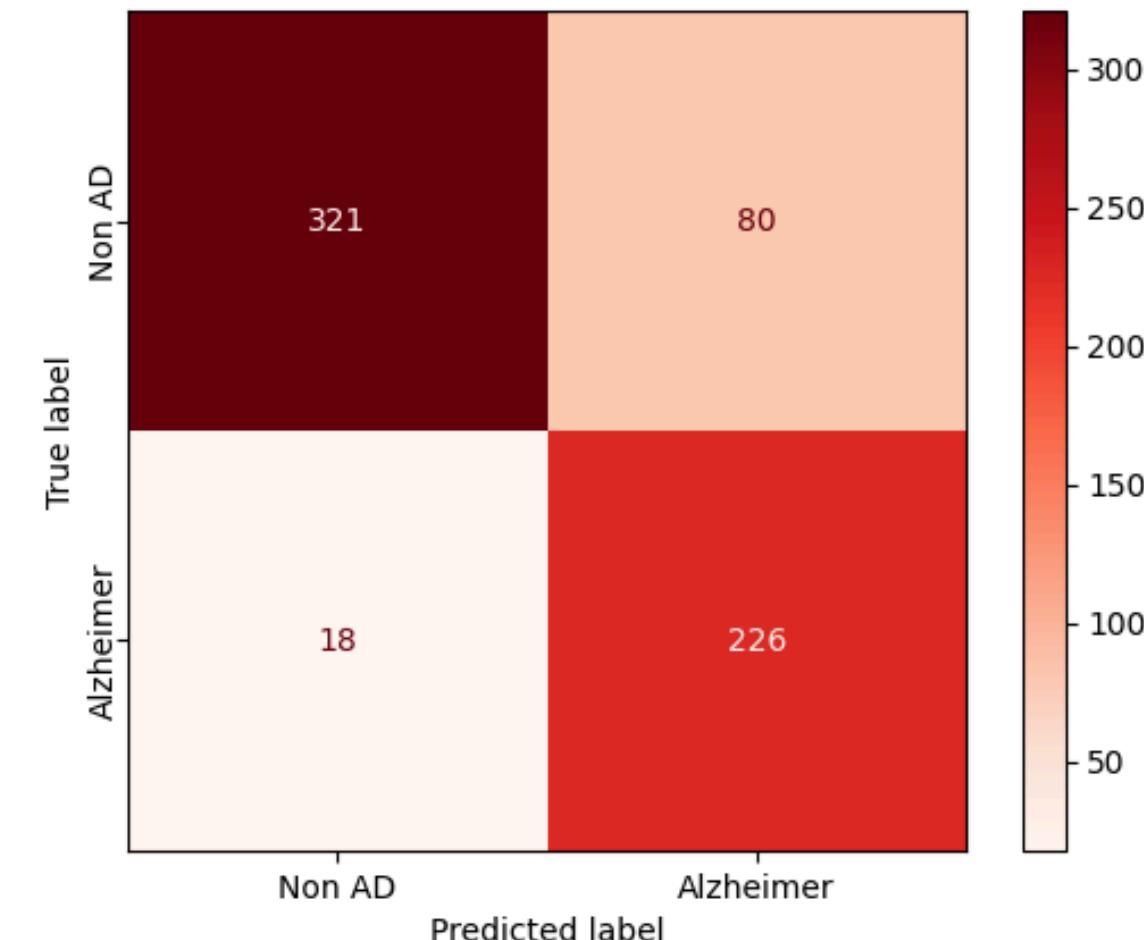
	Precision	Recall	F1-Score	ROC AUC
Hyperparameter Tuning LighGBM (Train)	0.952756	0.937984	0.945312	0.993252
Hyperparameter Tuning LighGBM (Test)	0.948718	0.909836	0.928870	0.946512
Hyperparameter Tuning LighGBM menggunakan SMOTEENN (Train)	0.811130	0.932171	0.867448	0.959905
Hyperparameter Tuning LighGBM menggunakan SMOTEENN (Test)	0.792982	0.926230	0.854442	0.933435

	Precision	Recall	F1-Score	ROC AUC
Hyperparameter Tuning XG-Boost (Train)	0.972112	0.945736	0.958743	0.995226
Hyperparameter Tuning XG-Boost (Test)	0.943966	0.897541	0.920168	0.944432
Hyperparameter Tuning XG-Boost menggunakan SMOTEENN (Train)	0.715758	0.941860	0.813389	0.940373
Hyperparameter Tuning XG-Boost menggunakan SMOTEENN (Test)	0.738562	0.926230	0.821818	0.919464

Confusion Matrix of Tuned with SMOTEENN Model LightGBM

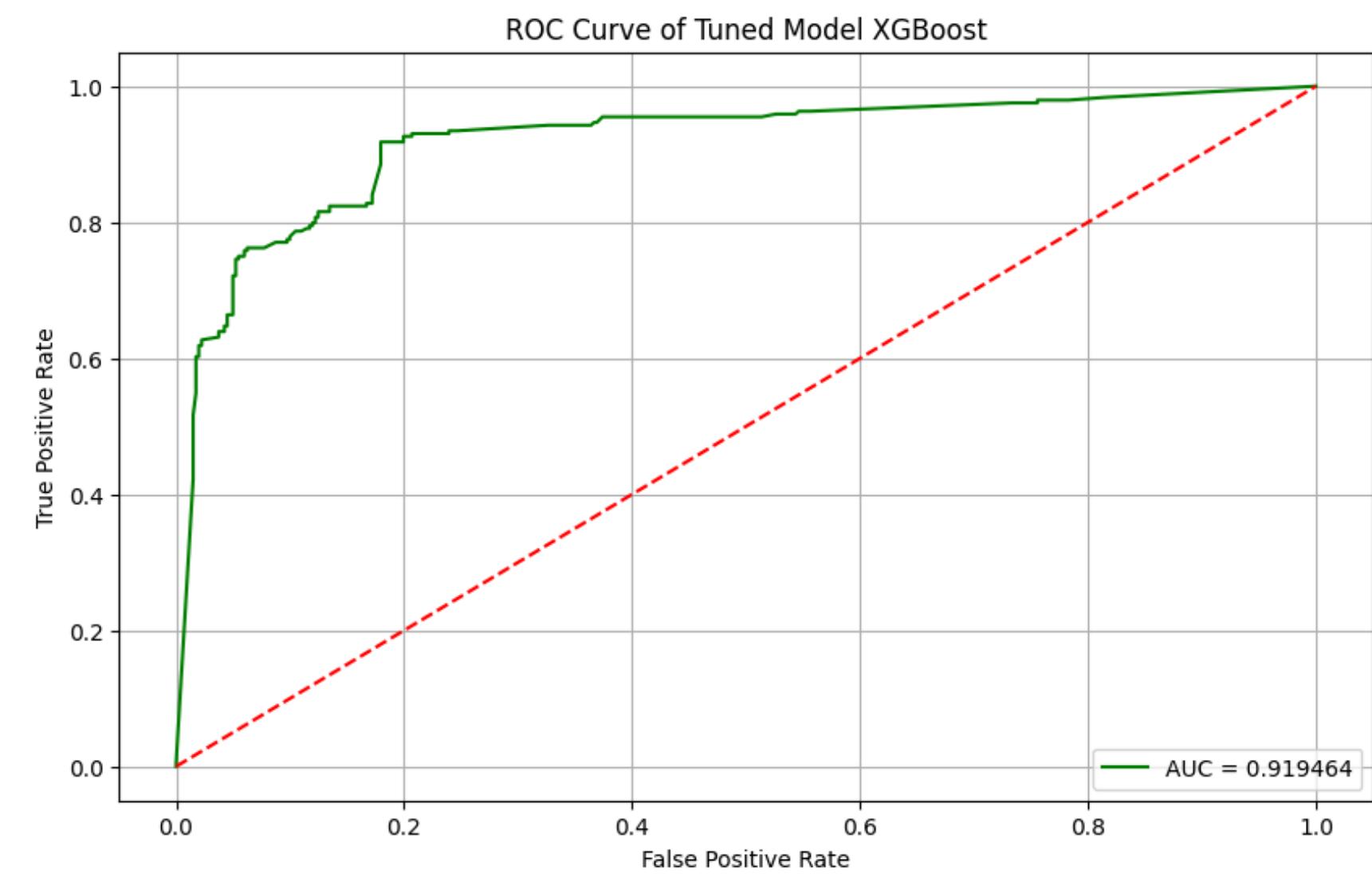
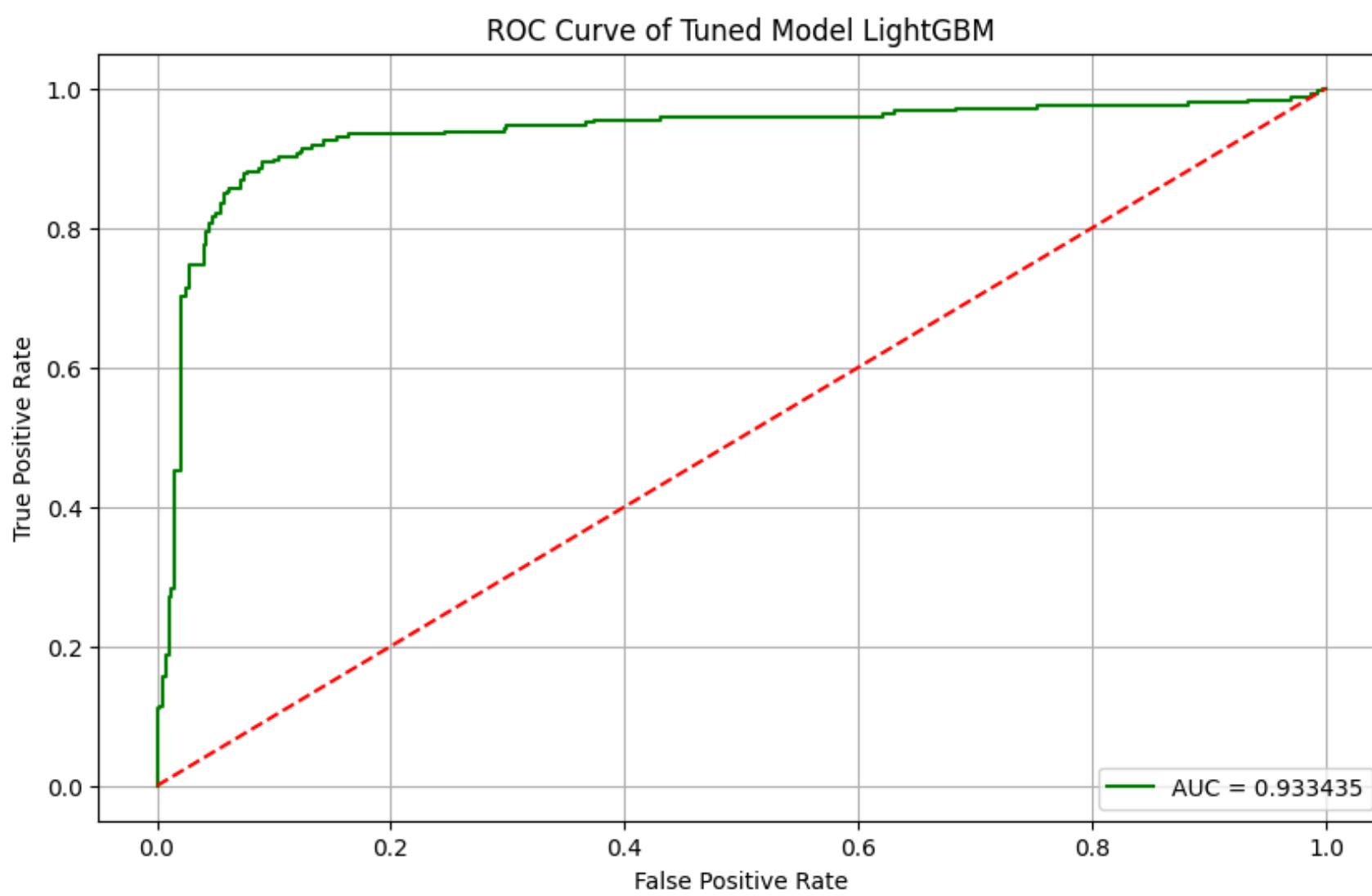


Confusion Matrix of Tuned with SMOTEENN Model XG-Boost



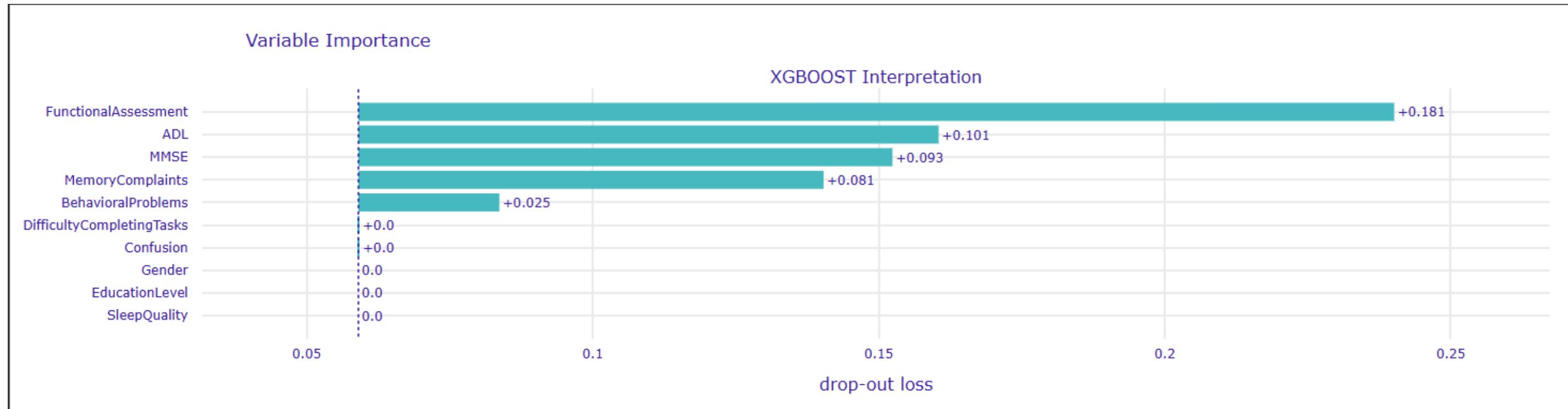
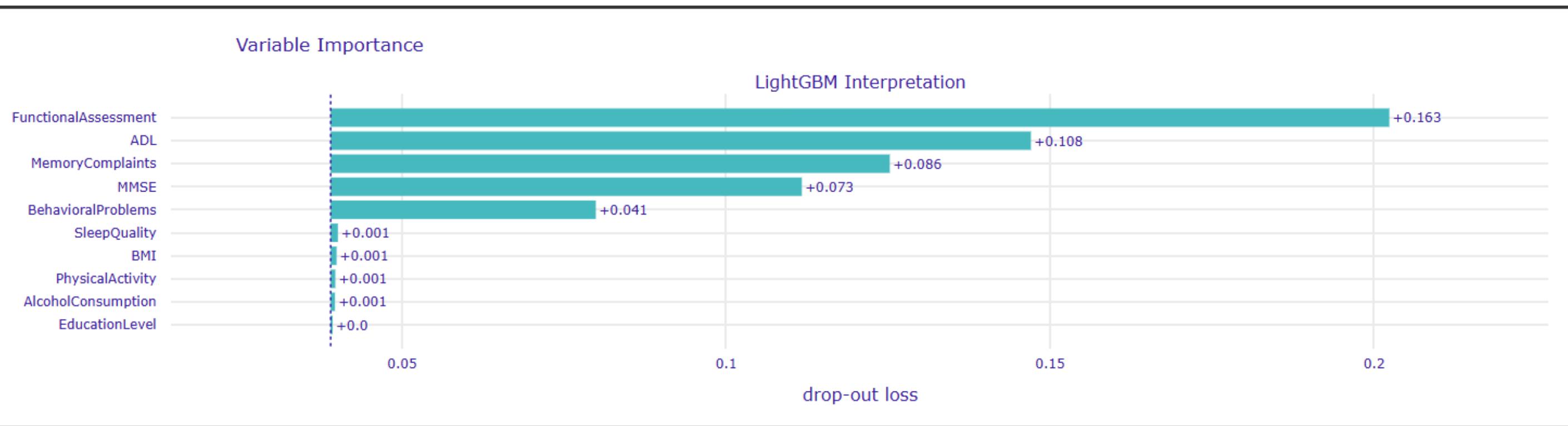
# Pemodelan Klasifikasi (1)

- ROC-AUC Curve



# Pemodelan Klasifikasi (1)

- Feature Importance



## Pemodelan Klasifikasi (2)

- Rasio split data 70 : 30
- Menggunakan Min Max Scaler
- Hanya membuang kolom 'DoctorInCharge', 'PatientID'

	Model	Precision (train)	Recall (train)	F1-Score (train)	ROC AUC (train)	Precision (test)	Recall (test)	F1-Score (test)	ROC AUC (test)
0	Logistic Regression	0.731219	0.848837	0.785650	0.916304	0.720930	0.762295	0.741036	0.880320
1	Decision Tree	1.000000	1.000000	1.000000	1.000000	0.850622	0.840164	0.845361	0.875194
2	Random Forest	1.000000	1.000000	1.000000	1.000000	0.953608	0.758197	0.844749	0.944217
3	XGBoost	1.000000	1.000000	1.000000	1.000000	0.947826	0.893443	0.919831	0.946680
4	LightGBM	1.000000	1.000000	1.000000	1.000000	0.951111	0.877049	0.912580	0.947243
5	KNN	0.824022	0.571705	0.675057	0.882882	0.629139	0.389344	0.481013	0.711086
6	SVM	0.852837	0.932171	0.890741	0.968703	0.739300	0.778689	0.758483	0.879420
7	MLP	0.790164	0.467054	0.587089	0.861428	0.758865	0.438525	0.555844	0.814807

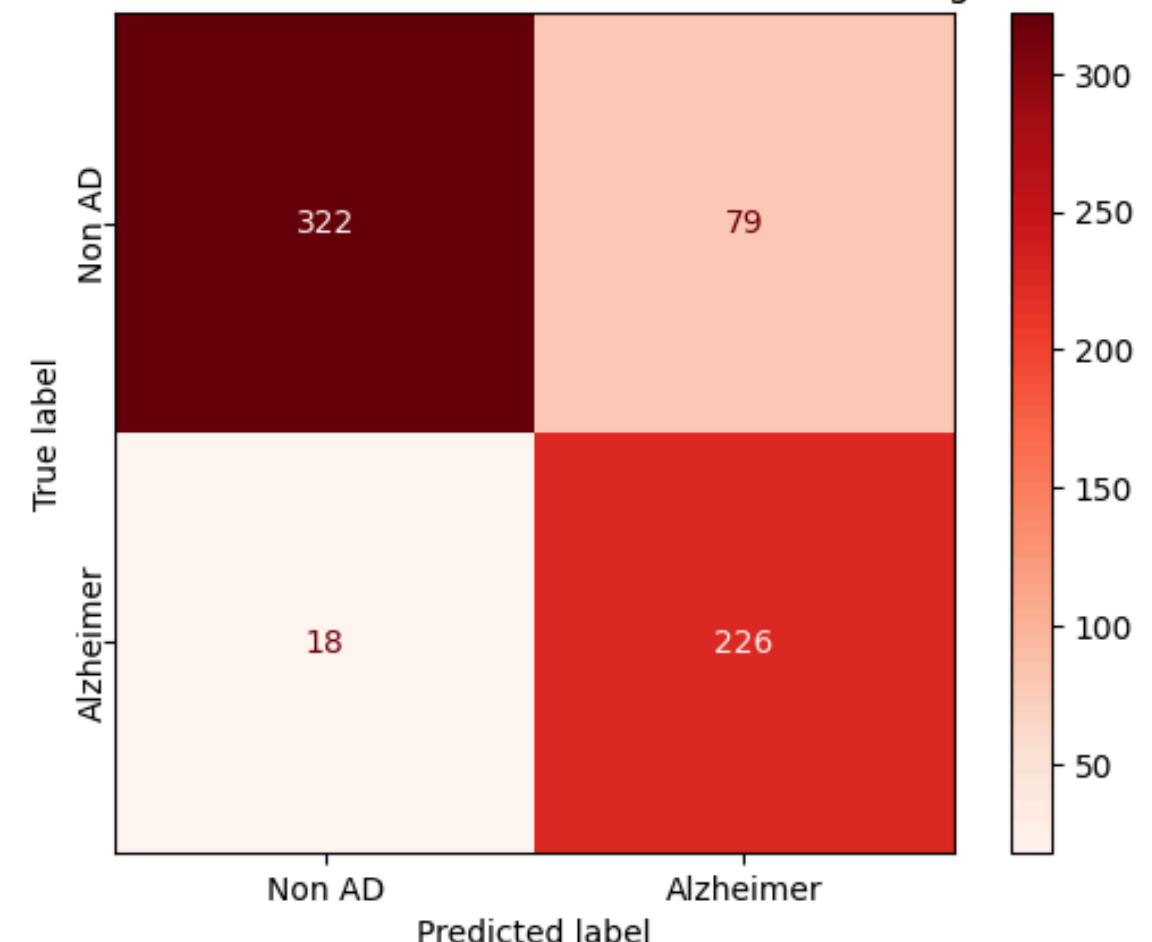
## Pemodelan Klasifikasi (2)

- Hyperparameter Tuning
- Hyperparameter Tuning with SMOTEENN

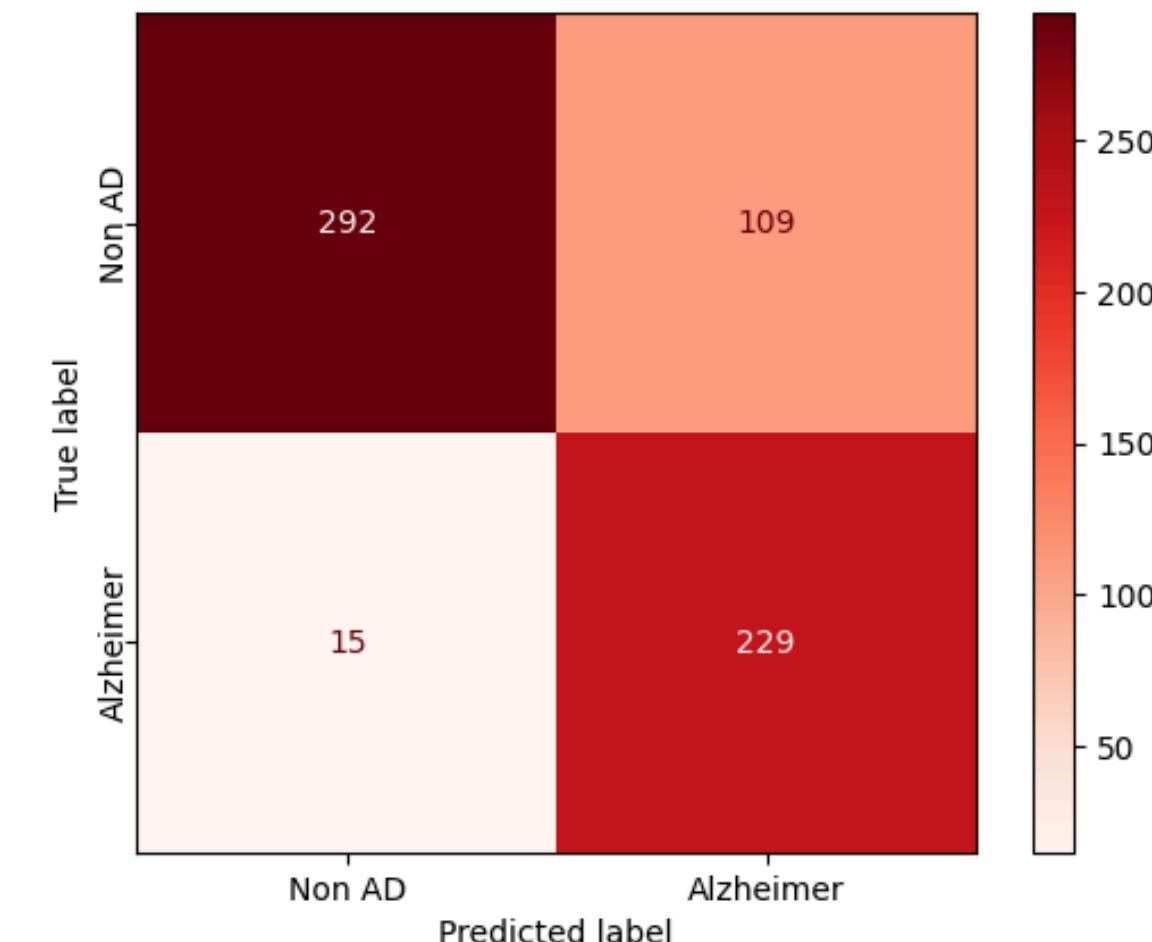
	Precision	Recall	F1-Score	ROC AUC
Hyperparameter Tuning LighGBM (Train)	0.949119	0.939922	0.944499	0.976311
Hyperparameter Tuning LighGBM (Test)	0.953191	0.918033	0.935282	0.946578
Hyperparameter Tuning LighGBM menggunakan SMOTEENN (Train)	0.734121	0.963178	0.833194	0.960003
Hyperparameter Tuning LighGBM menggunakan SMOTEENN (Test)	0.740984	0.926230	0.823315	0.924318

	Precision	Recall	F1-Score	ROC AUC
Hyperparameter Tuning XG-Boost (Train)	1.000000	1.000000	1.000000	1.000000
Hyperparameter Tuning XG-Boost (Test)	0.946903	0.877049	0.910638	0.944279
Hyperparameter Tuning XG-Boost menggunakan SMOTEENN (Train)	0.632522	0.957364	0.761758	0.940255
Hyperparameter Tuning XG-Boost menggunakan SMOTEENN (Test)	0.677515	0.938525	0.786942	0.918733

Confusion Matrix of Tuned with SMOTEENN Model LightGBM

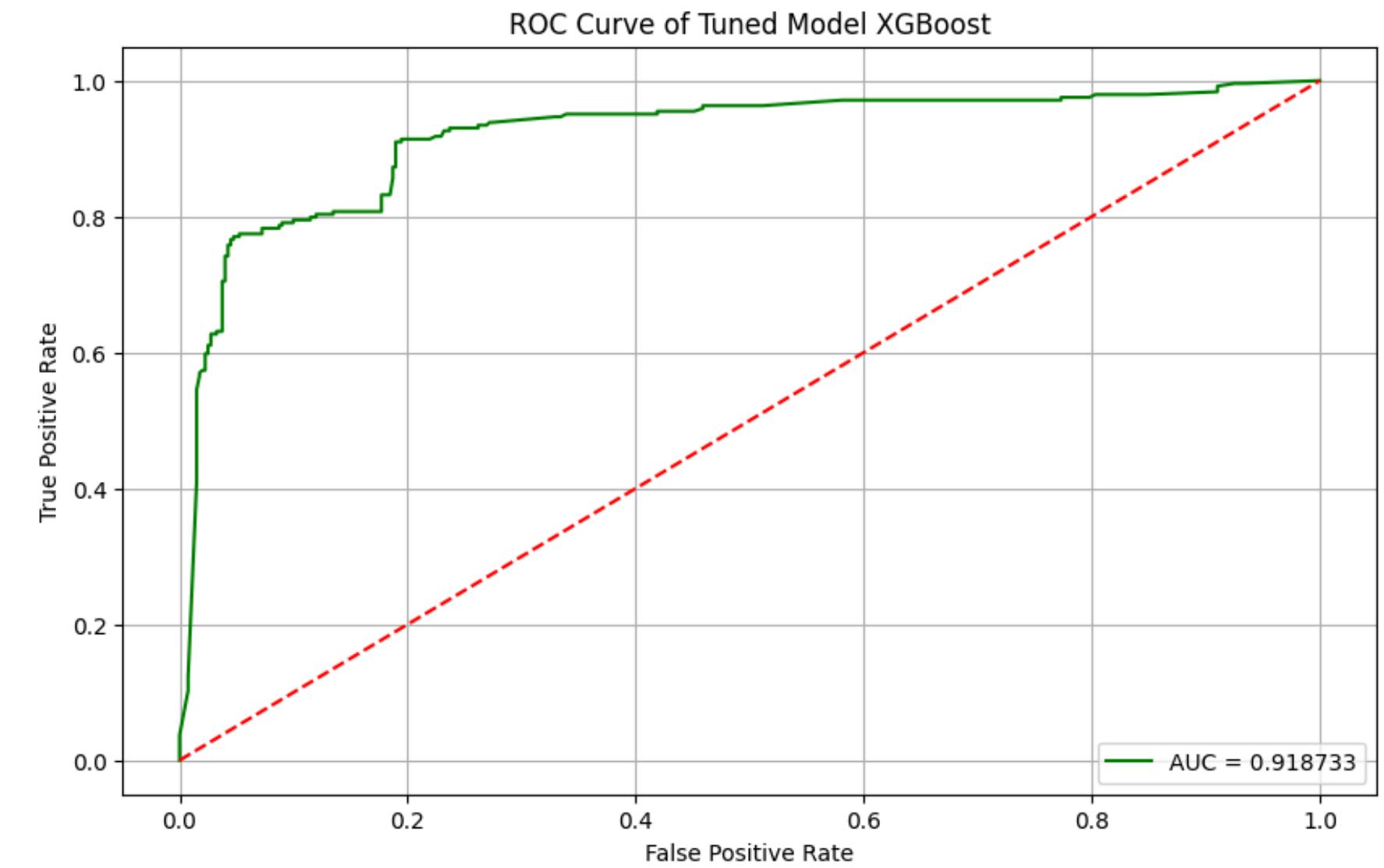
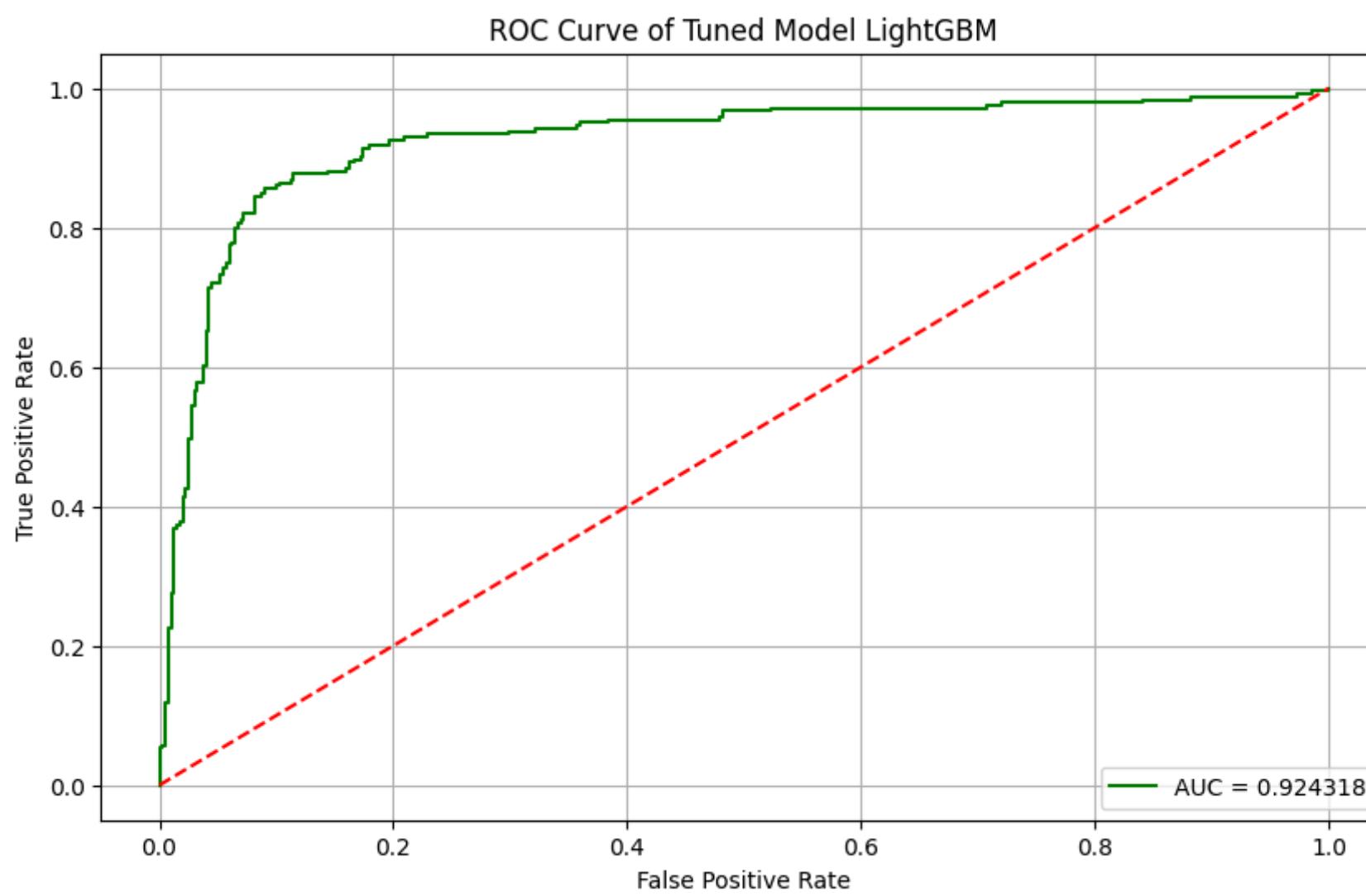


Confusion Matrix of Tuned with SMOTEENN Model XG-Boost



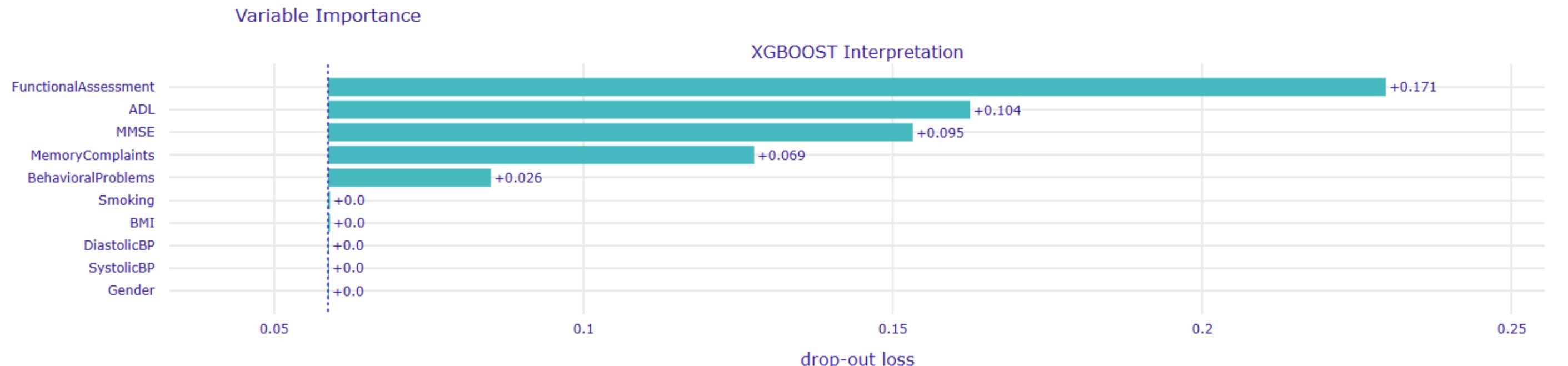
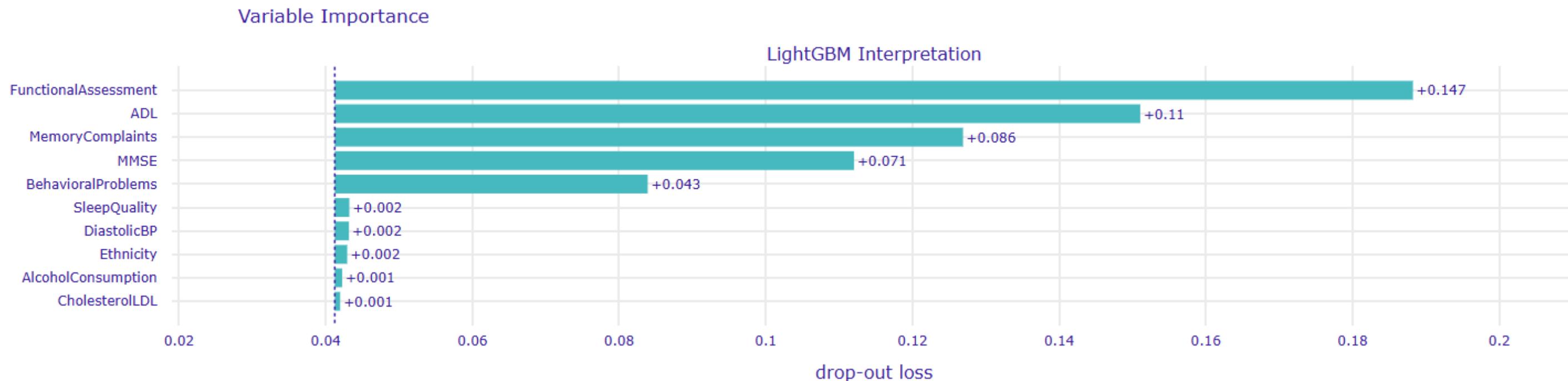
# Pemodelan Klasifikasi (2)

- ROC-AUC Curve



# Pemodelan Klasifikasi (2)

- Feature Importance



STUDIO SHODWE

# Thank You!

*Presented By: Rufus Stewart*

+123-456-7890

[www.reallygreatsite.com](http://www.reallygreatsite.com)

@reallygreatsite

