



PREDIKSI DEPRESI DENGAN MENGEKSPLORASI DATA KESEHATAN MENTAL MENGGUNAKAN REGRESI LOGISTIK

Authors:

Bryan Evans Simamora
Agnes Monica Sanjani Harefa
Yohana Christine Sitanggang

Advisor :

Humasak Tommy Argo Simanjuntak, ST, M.ISD

1

Latar Belakang

Kesehatan mental merupakan salah satu aspek penting dalam mewujudkan kualitas hidup yang optimal, baik secara individu maupun dalam konteks sosial dan profesional. Salah satu gangguan kesehatan mental yang umum terjadi adalah depresi, yang dapat berdampak signifikan terhadap produktivitas kerja, pengambilan keputusan, serta relasi sosial individu. World Health Organization (WHO) menyatakan bahwa masalah kesehatan mental, termasuk depresi, sering kali tidak memperoleh perhatian yang memadai, khususnya di lingkungan kerja, meskipun dampaknya sangat merugikan secara psikologis maupun ekonomi. Berbagai faktor seperti usia, tekanan akademik, tekanan pekerjaan, kualitas tidur, serta stres finansial diketahui berkontribusi terhadap risiko terjadinya depresi. Melalui pendekatan berbasis data, penelitian ini bertujuan untuk menganalisis data survei sintetis menggunakan algoritma Logistic Regression guna membangun model prediksi risiko depresi. Model ini diharapkan dapat membantu dalam proses identifikasi dini terhadap individu yang berisiko mengalami depresi, sehingga dapat menjadi dasar bagi pengambilan kebijakan preventif dan intervensi yang lebih tepat sasaran, khususnya di lingkungan kerja atau institusi pendidikan.

2

Simulator Website

Check Depression Prediction

Gender

Male

Age

City (A=1, B=2, ...)

Working Status

Working Professional

Profession (A=1, ... Z=26)

Degree (numeric)

Academic Pressure

Work Pressure

Study Satisfaction

Job Satisfaction

Dietary Habits

Healthy

Suicidal Thoughts

No

Work/Study Hours

Financial Stress

Family History

No

Sleep Duration (hours)

Predict

Result: Not Depressed

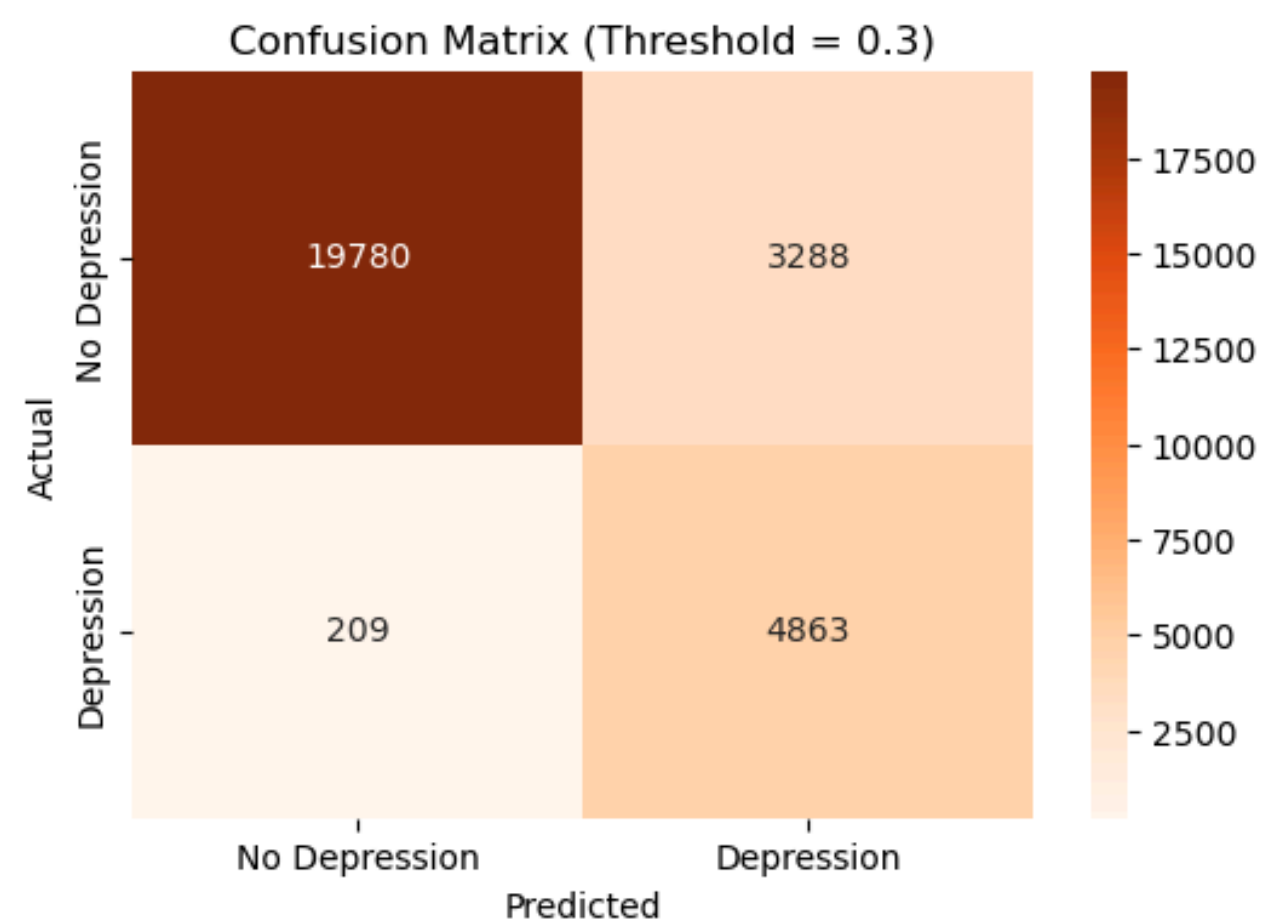
3

Tujuan Penelitian

1. Menganalisis faktor-faktor yang berkontribusi terhadap risiko depresi berdasarkan data survei kesehatan mental sintetis.
2. Membangun model klasifikasi menggunakan algoritma Logistic Regression untuk memprediksi kemungkinan seseorang mengalami depresi.
3. Mengevaluasi performa model prediktif menggunakan metrik evaluasi seperti akurasi, precision, recall, dan F1-score.

4

Confusion Matrix



5

Tahapan Penelitian

1. Business Understanding
Menentukan tujuan proyek memprediksi risiko depresi berdasarkan data survei.
2. Data Understanding
Mengumpulkan, mendeskripsikan, dan memvalidasi data (cek missing value, outlier, duplikasi).
3. Data Preparation
Seleksi, pembersihan, konstruksi, pelabelan, dan integrasi data agar siap diproses.
4. Modeling
Membangun model klasifikasi menggunakan Logistic Regression untuk prediksi depresi.
5. Evaluation
Mengevaluasi model menggunakan metrik seperti accuracy, precision, recall, dan F1-score.
6. Deployment
Men-deploy model agar bisa digunakan untuk prediksi secara real-time.

6

Kesimpulan

1. Algoritma Logistic Regression berhasil membangun model klasifikasi untuk memprediksi risiko depresi dengan tingkat akurasi sebesar 94%, yang menunjukkan bahwa metode ini efektif untuk kasus klasifikasi biner pada data kesehatan mental.
2. Faktor-faktor seperti usia, tekanan akademik, tekanan kerja, stres finansial, dan durasi tidur memiliki kontribusi signifikan terhadap kemungkinan seseorang mengalami depresi.
3. Meskipun dataset yang digunakan bersifat sintetis dan memiliki ketidakseimbangan kelas, model tetap menunjukkan performa yang baik, khususnya dalam mengidentifikasi individu yang berisiko mengalami depresi.
4. Hasil dari penelitian ini menunjukkan bahwa pendekatan berbasis data dapat digunakan sebagai dasar pengambilan keputusan dalam merancang intervensi kesehatan mental yang lebih efektif.