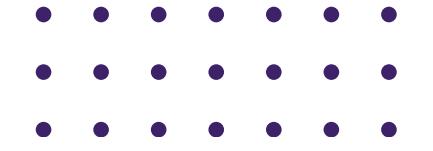




IFEST 2024



FINAL DATA ANALYSIS COMPETITION

AI IMPACT CLASSIFICATION IN SOUTHEAST ASIAN COUNTRIES: A MACHINE LEARNING APPROACH

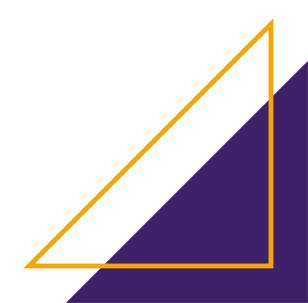
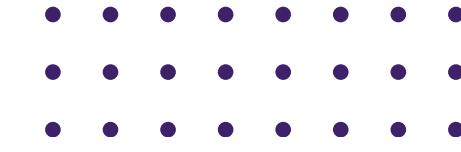
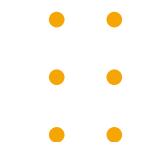
Tim Sudah Saatnya

Erlin Shofiana

Aulia Mirfah Setyo Ayu Damayanti

Fitri Hartanti





OVERVIEW

01 Latar Belakang dan Urgensi

02 Alur Pengerjaan

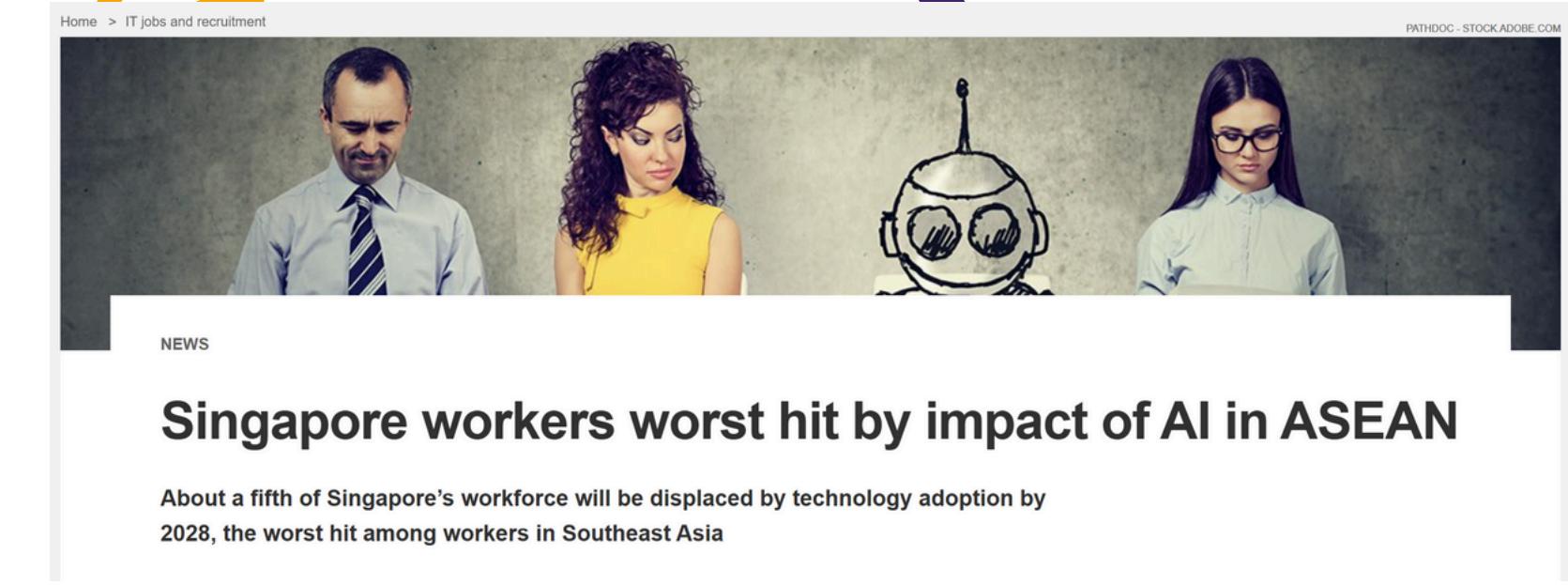
03 Visualisasi Data

04 Pembentukan Model

05 Kesimpulan dan Saran

LATAR BELAKANG DAN URGENSI

- Dikutip dari *The ASEAN Post*, adopsi AI memungkinkan enam ekonomi terbesar ASEAN mengurangi 28 juta pekerja pada 2028, dengan sektor pertanian diproyeksikan kehilangan 6,6 juta pekerjaan.
- Negara seperti Filipina, Thailand, dan Indonesia berisiko tinggi di sektor *Bussiness Process Outsourcing* (BPO) dan *contact centers*.
- • • Studi oleh *Cisco and Oxford Economics* menunjukkan bahwa Singapura menghadapi risiko tertinggi di ASEAN, dengan sekitar 21% pekerja penuh waktu berpotensi kehilangan pekerjaan akibat adopsi AI.
- Klasifikasi pekerjaan berdasarkan dampak AI ("less," "medium," "high") sangat penting untuk memetakan sektor yang paling terancam oleh otomatisasi.



Home > IT jobs and recruitment

NEWS

Singapore workers worst hit by impact of AI in ASEAN

About a fifth of Singapore's workforce will be displaced by technology adoption by 2028, the worst hit among workers in Southeast Asia

CNA Top Stories Latest News Asia East Asia Singapore Commentary Sustainability CNA Insider Lifestyle Watch Listen + All Sections

AI berkembang pesat di Asia Tenggara, apakah pekerjaan manusia akan terancam?

Dari rumah sakit hingga layanan call centre, AI menimbulkan kecemasan sekaligus harapan terhadap peluang kerja di masa depan.

28 million ASEAN jobs threatened by AI over the next decade

Thaiger · Published: 11:01, 13 September 2018 | Updated: 11:05, 21 December 2023

Share this article Facebook Twitter LinkedIn Messenger Share via Email Print

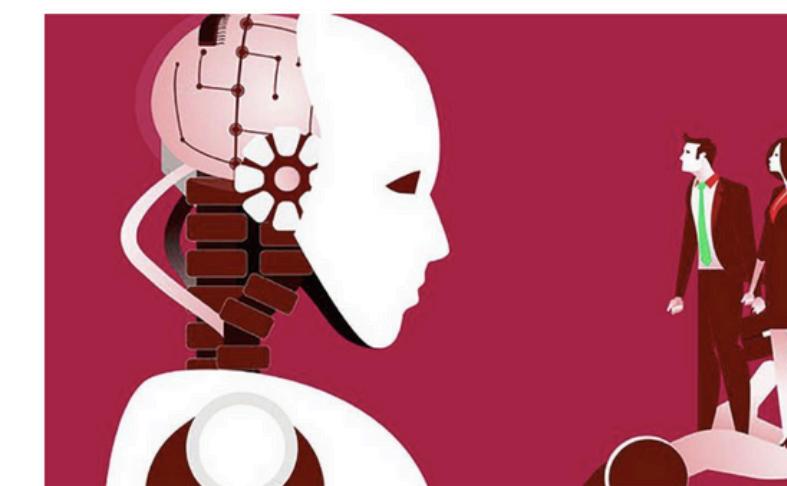
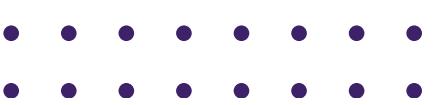


PHOTO: Krungsri Finnovate

AI threatens ASEAN jobs

Stephen Chin
20 September 2018





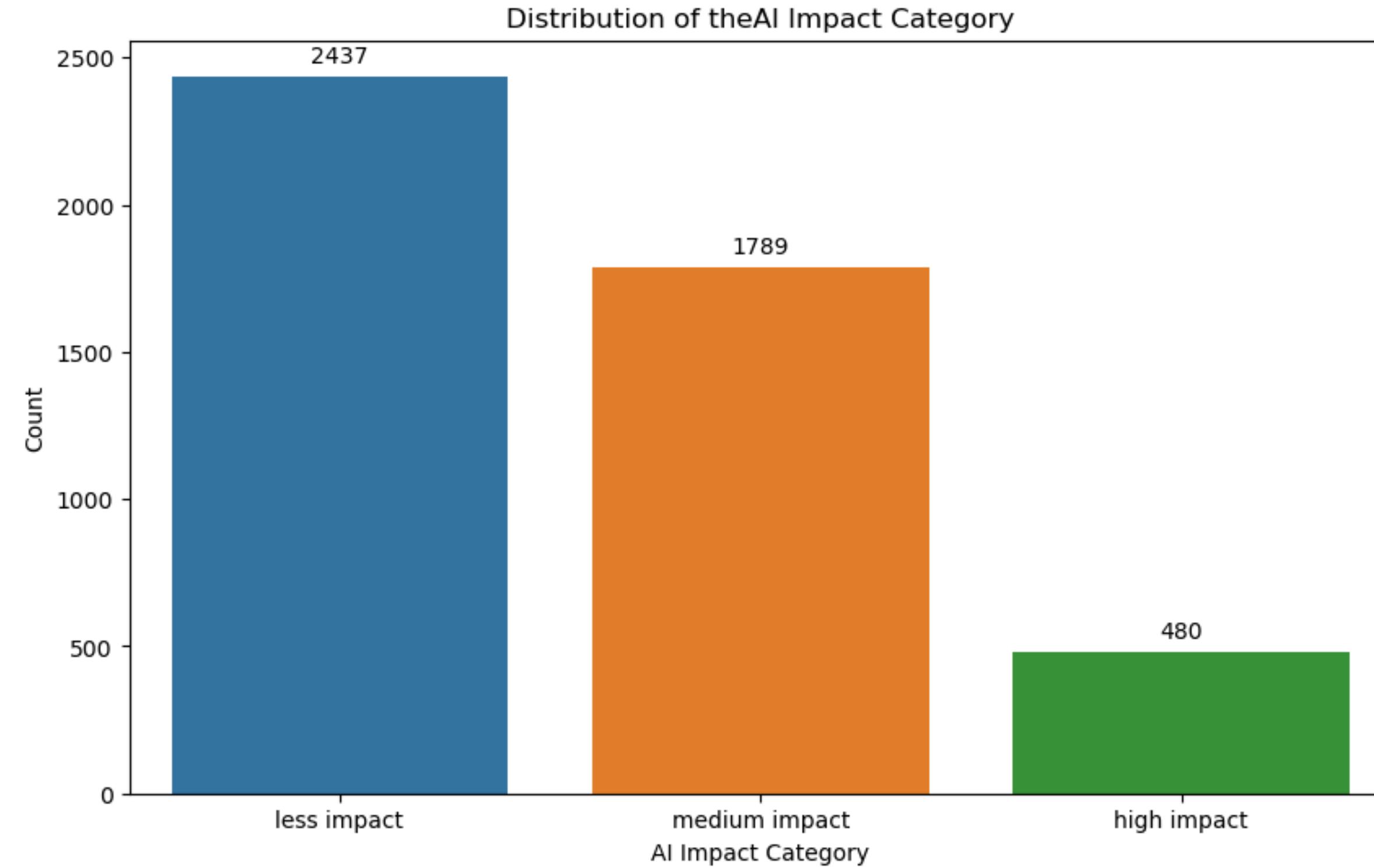
VARIABEL TARGET

'AI Impact Category'

Menentukan variabel target berdasarkan kategorisasi variabel 'AI Impact' dengan pembagian sebagai berikut:

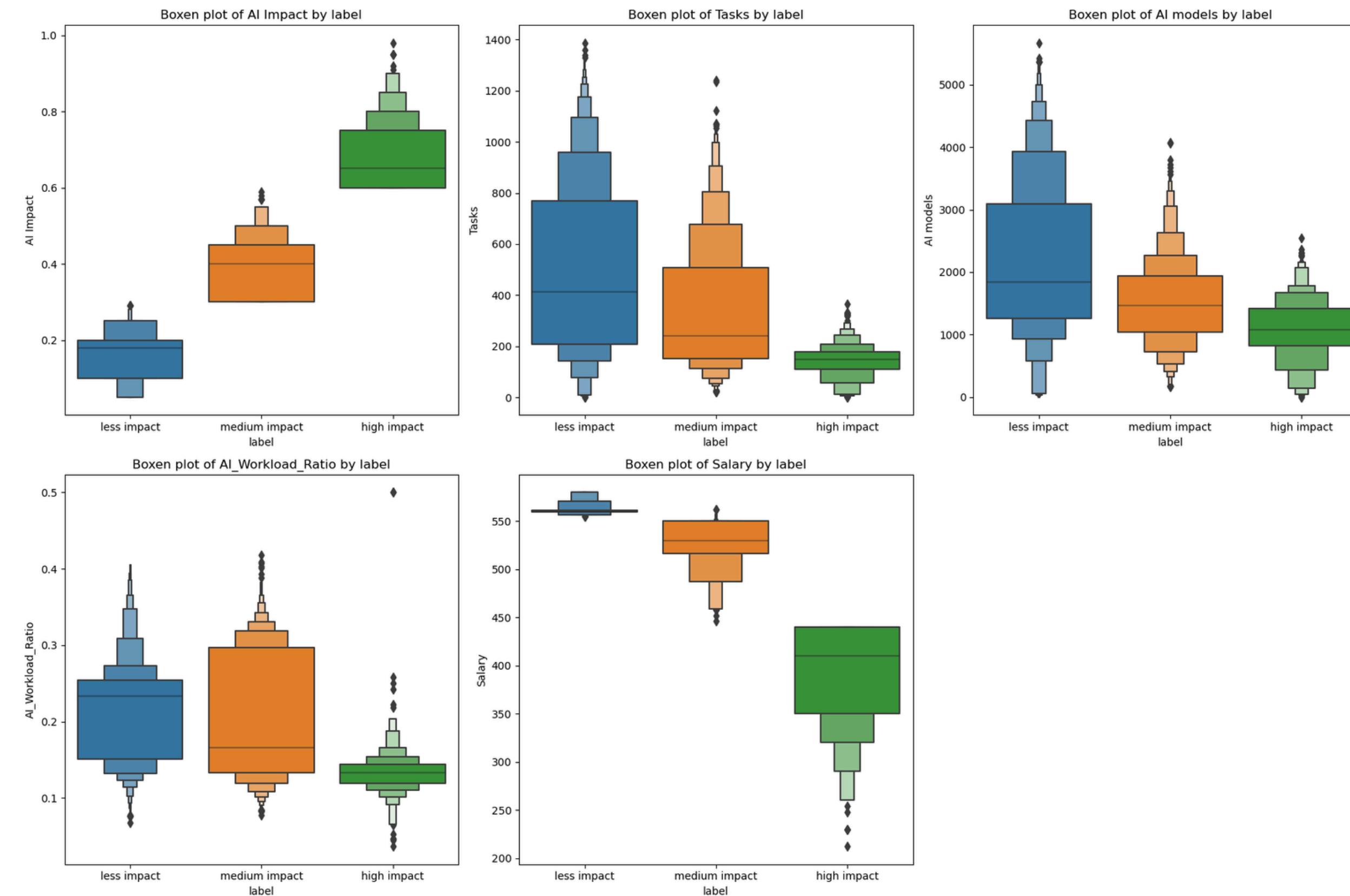
- $0 \leq \text{AI Impact} < 30\% : \text{Dampak rendah}$
- $30 \leq \text{AI Impact} < 60\% : \text{Dampak sedang}$
- $\text{AI Impact} \geq 60\% : \text{Dampak tinggi}$

VISUALISASI DATA



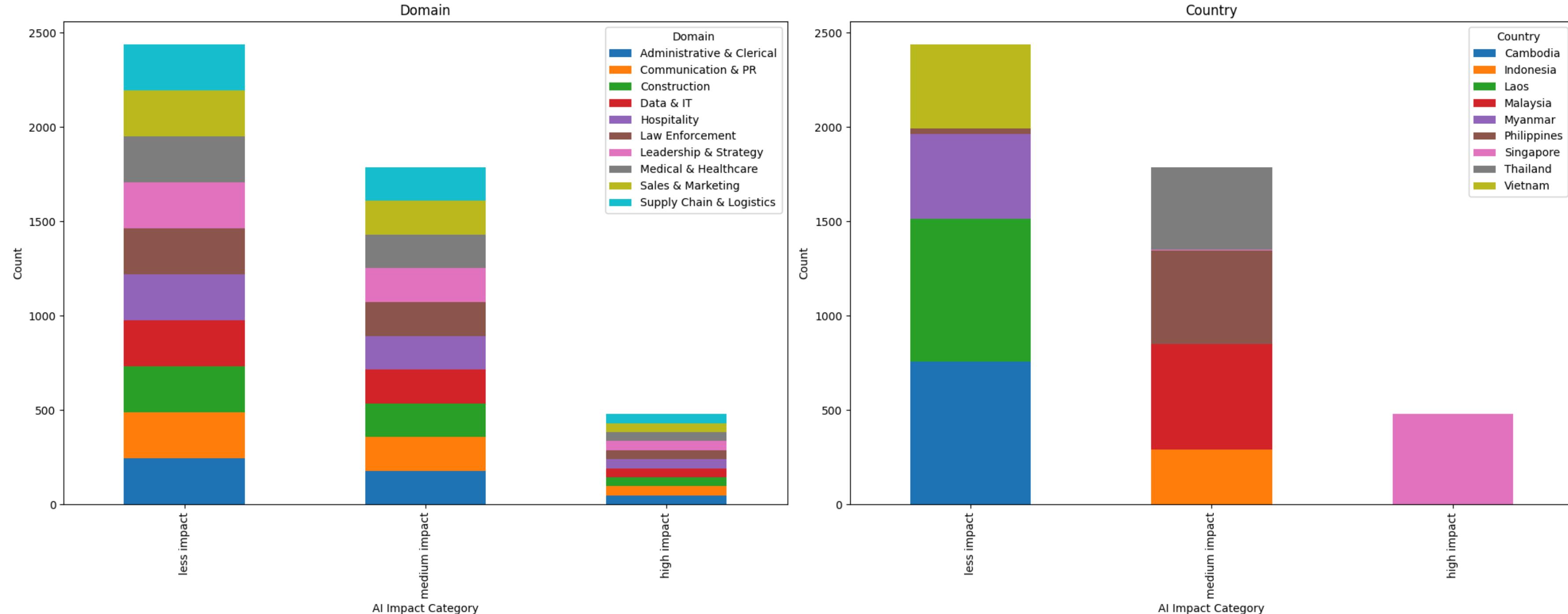
SEBARAN TARGET (AI IMPACT CATEGORY)

VISUALISASI DATA



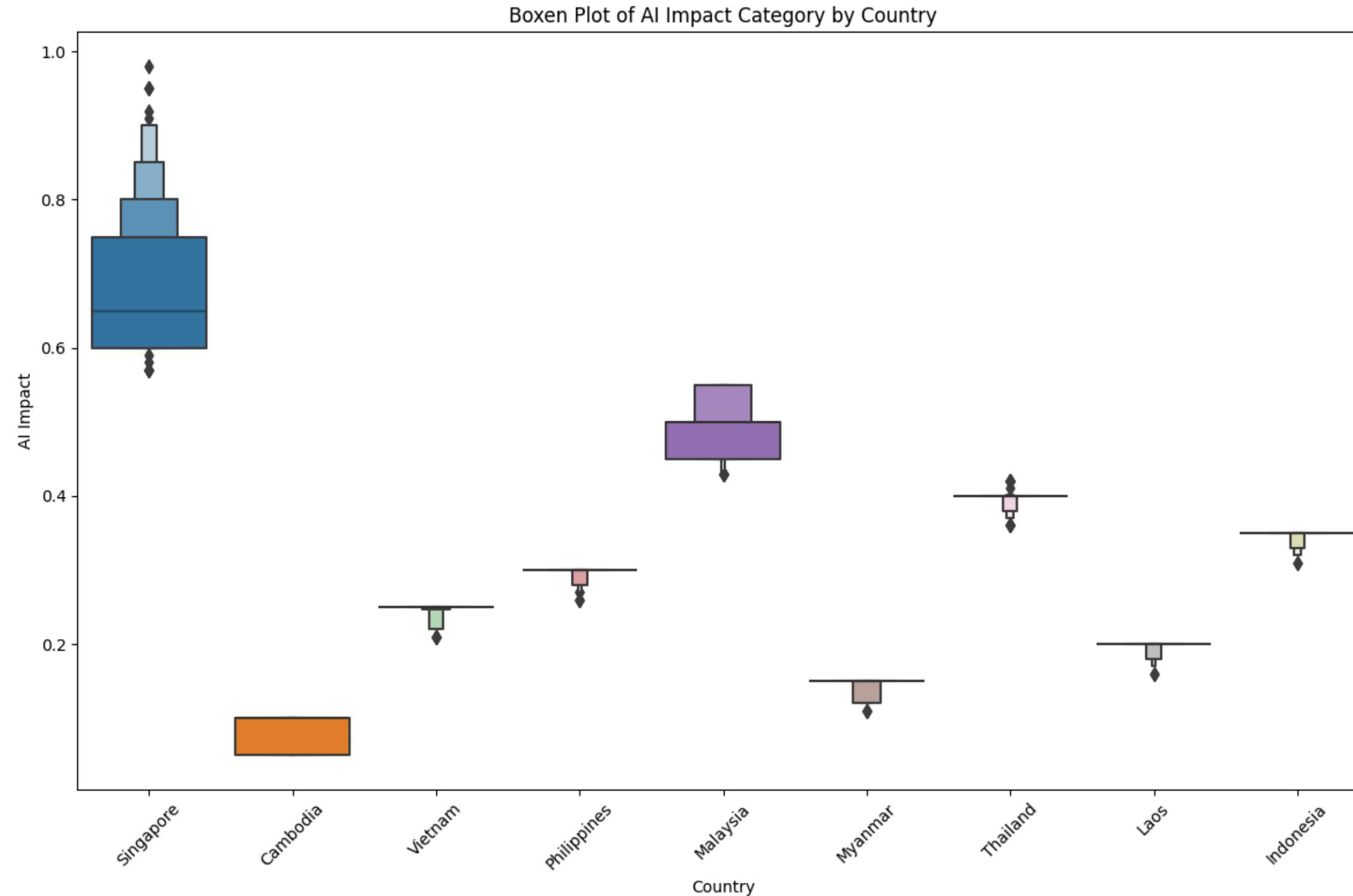
SEBARAN VARIABEL NUMERIK

VISUALISASI DATA



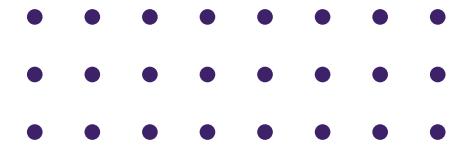
SEBARAN VARIABEL KATEGORIK

VISUALISASI DATA



SEBARAN VARIABEL KATEGORIK

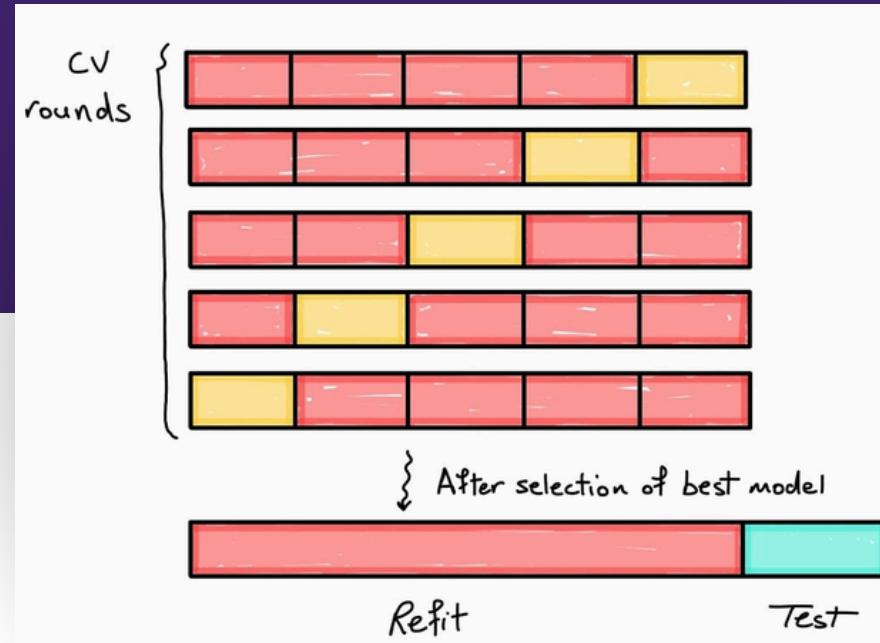
PRA-PEMROSESAN DATA



- 1** Menghapus kolom 'Job Titles'
- 2** Mengubah nilai infinity pada variabel 'AI_Workload_Ratio' dengan NaN. Kemudian mengisinya dengan rata rata 'AI_Workload_Ratio' untuk kelas less_impact
- 3** Mengonversi kategori variabel 'AI Impact Category' ke dalam bentuk numerik.
 - 'less impact' dikonversi menjadi 0,
 - 'medium impact' menjadi 1,
 - dan 'high impact' menjadi 2.
- 4** Menghitung median gaji per negara kemudian gaji setiap job title dikategorikan sebagai 'Above Median' atau 'Below Median' berdasarkan apakah gaji mereka lebih besar atau lebih kecil dari median di negara tersebut.



PEMBENTUKAN MODEL



	Model	Cross-Val Train Accuracy
1	XGBoost	0.998406
2	Extra Tree	0.997343
6	Decision Tree	0.997343
4	Naive Bayes	0.992828
5	AdaBoost	0.992828
3	K-Nearest Neighbors	0.660469
0	SVM	0.560310

	precision	recall	f1-score	support
0	1.000000	1.000000	1.000000	488
1	0.997207	1.000000	0.998601	357
2	1.000000	0.989691	0.994819	97
accuracy				0.998938
macro avg	0.999069	0.996564	0.997807	942
weighted avg	0.998941	0.998938	0.998936	942

Splitting Data

Data dibagi **80% untuk pelatihan** dan **20% untuk pengujian**, kemudian data pelatihan di-**cross-validate** dengan **5-fold** untuk memastikan evaluasi model lebih akurat.

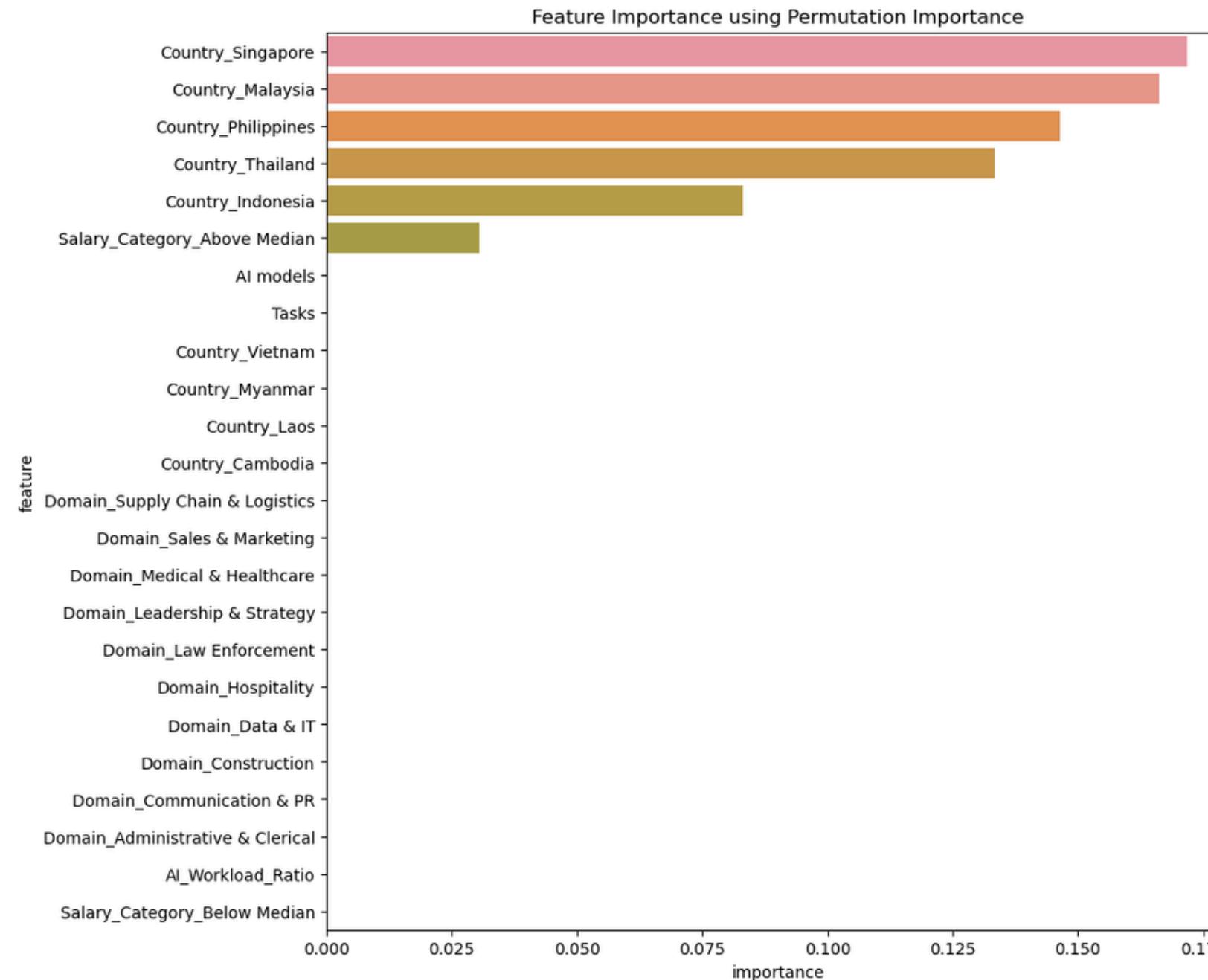
Pembentukan Model

Hasil pembentukan tiap model dengan metrik evaluasi akurasi menghasilkan model dengan akurasi tertinggi yaitu **XGBoost**.

Pengujian Model

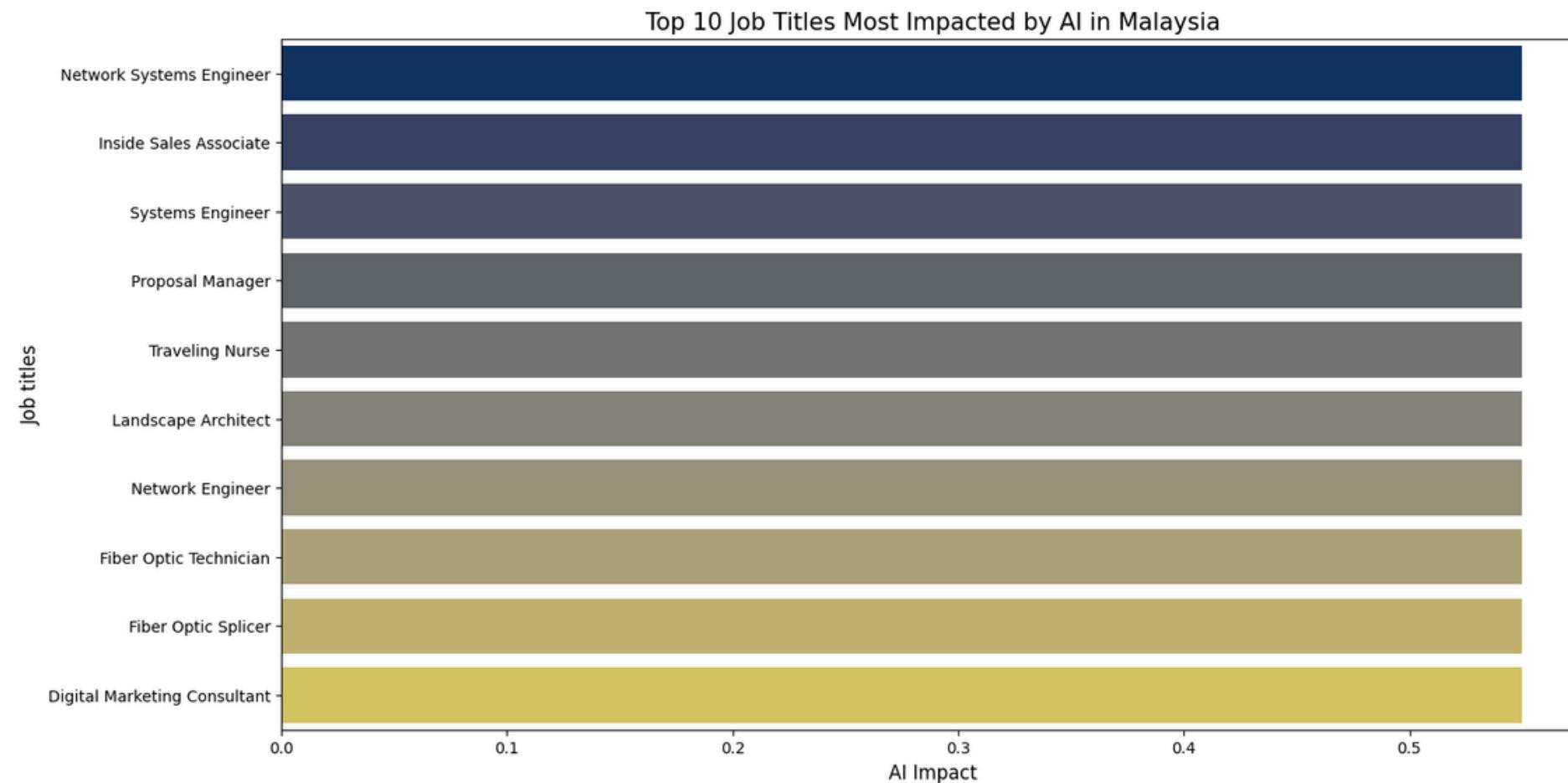
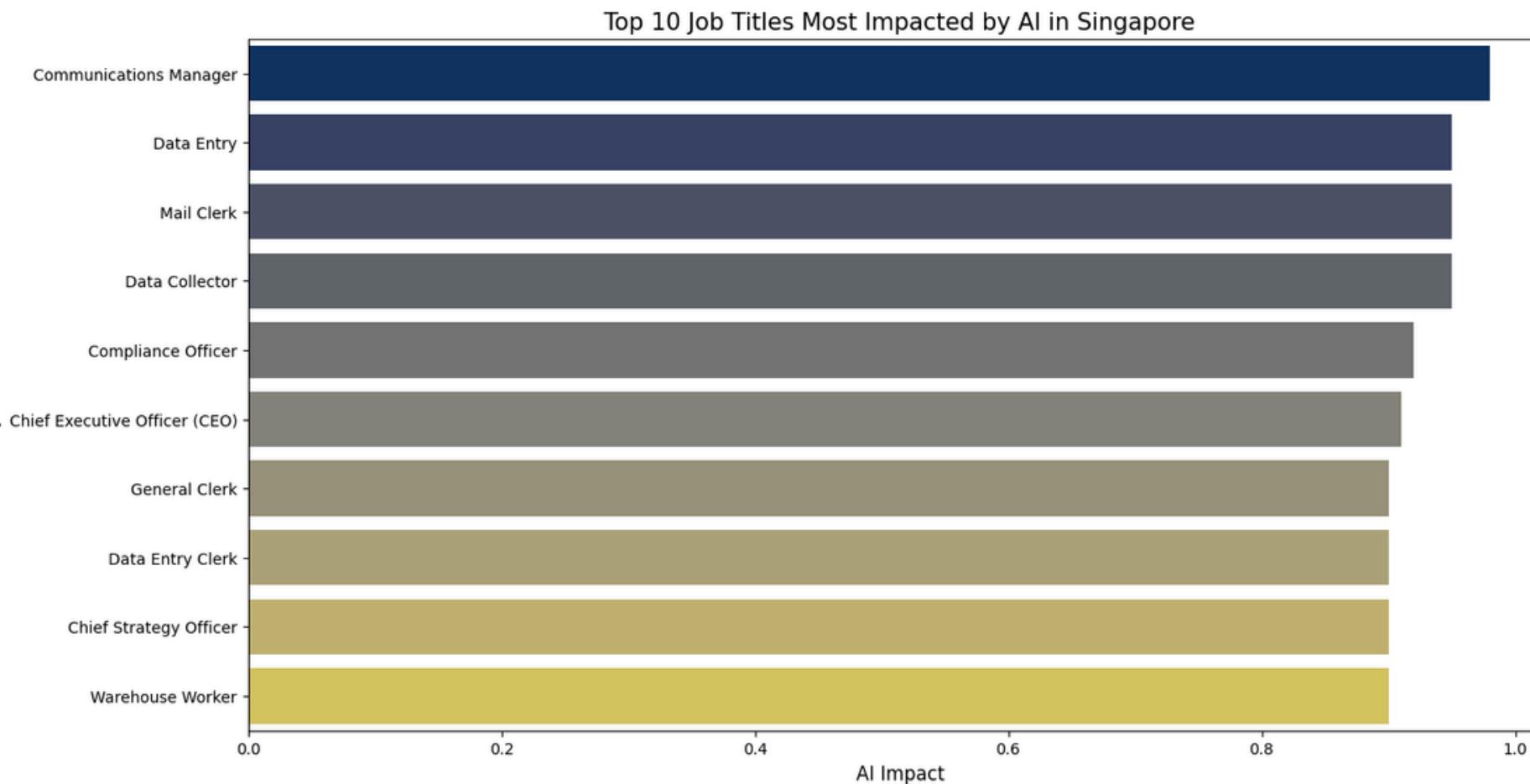
Hasil *fitting model* terbaik (XGBoost) pada data test dengan akurasi sebesar **99,89%**.

FEATURE IMPORTANCE



Untuk mengidentifikasi fitur mana yang memiliki dampak terbesar pada prediksi model, dibentuk *feature importance* berupa *permutation importance* dari model terbaik yaitu dengan algoritma XGBoost. Berdasarkan hasil *feature importance*, fitur yang memiliki pengaruh paling tinggi terhadap prediksi model adalah Negara Singapura, Malaysia, Filipina, Thailand, dan Indonesia, serta kelompok gaji. Hal ini menunjukkan bahwa fitur-fitur tersebut memberikan kontribusi terbesar dalam menentukan probabilitas *AI Impact* apakah tinggi, sedang, atau rendah.

TOP 10 PEKERJAAN PALING TERDAMPAK OLEH AI DI SINGAPURA DAN MALAYSIA



KESIMPULAN DAN SARAN

Kesimpulan

Setelah melalui tahapan pra-pemrosesan data, model XGBoost dipilih sebagai yang terbaik dengan akurasi 99,89%. Analisis feature importance menunjukkan bahwa fitur Negara (Singapura, Malaysia, Filipina, Thailand, Indonesia) dan kelompok gaji memiliki pengaruh terbesar dalam menentukan AI Impact. Model ini berfungsi sebagai alat prediksi untuk menilai dampak AI pada berbagai pekerjaan berdasarkan fitur-fitur tersebut.

Saran

- Perlunya kebijakan mitigasi AI yang disesuaikan dengan negara dan kelompok pekerja yang rentan, terutama di sektor-sektor berisiko tinggi.
- Negara-negara ASEAN perlu fokus pada pekerja dengan gaji rendah melalui program peningkatan keterampilan dan pelatihan ulang untuk mengurangi dampak negatif AI terhadap tenaga kerja.



TERIMA KASIH