

Paradoks Penyerapan Tenaga Kerja: Prediksi Tingkat *Underemployment* di 34 Provinsi Indonesia

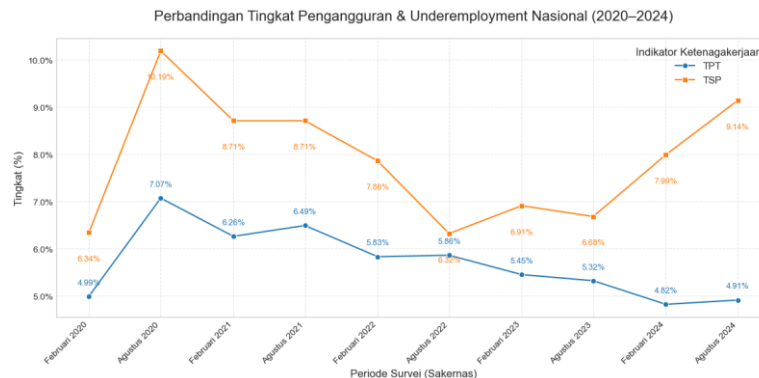
Najla Dhia Rusydi

164221043

najla.dhia.rusydi-2022@ftmm.unair.ac.id

A. Pendahuluan

Saat ini Indonesia tengah menghadapi sebuah paradoks dalam pasar tenaga kerja, di tengah pesatnya pertumbuhan ekonomi dan percepatan digitalisasi, kualitas penyerapan tenaga kerja belum menunjukkan perbaikan yang signifikan. Masalah utamanya bukanlah sekadar pengangguran terbuka, melainkan tingginya angka *underemployment*, yaitu kondisi di mana tenaga kerja bekerja kurang dari 35 jam per minggu namun masih menginginkan jam kerja tambahan. Fenomena ini menunjukkan bahwa banyak pekerjaan yang tersedia bersifat tidak penuh dan tidak memberikan jaminan kesejahteraan yang memadai.



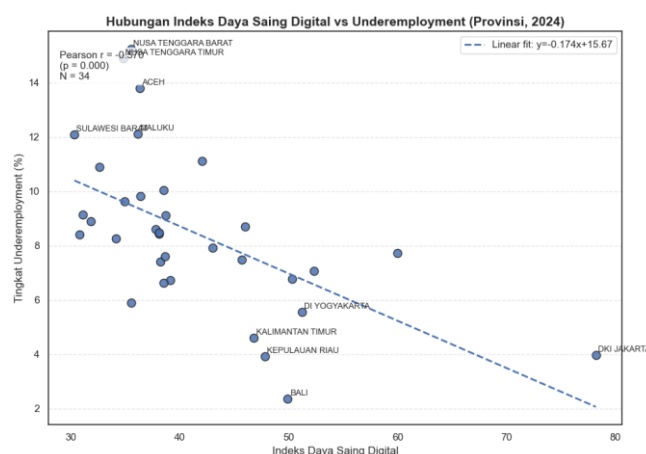
Gambar 1. Tren TPT dan *Underemployment* di Indonesia

Sumber: BPS (Sakernas, 2024).

Berdasarkan data Badan Pusat Statistik (BPS, Sakernas 2024), tingkat *underemployment* nasional mencapai 9,76% dari total angkatan kerja, jauh lebih tinggi dibandingkan tingkat pengangguran terbuka yang hanya 5,32%. Meskipun perekonomian nasional tumbuh positif, dengan peningkatan PDRB per kapita hingga sekitar Rp 75 juta pada tahun 2023, kualitas pekerjaan belum sepenuhnya mencerminkan hasil dari pertumbuhan ekonomi tersebut. Situasi ini memunculkan pertanyaan krusial, mengapa pertumbuhan ekonomi dan digitalisasi belum mampu secara efektif mengatasi masalah kualitas pekerjaan di 34 provinsi Indonesia?

Penelitian sebelumnya telah memberikan pemahaman awal mengenai faktor-faktor yang memengaruhi *underemployment* di Indonesia. Studi oleh Pratomo (2015) menemukan bahwa baik karakteristik individu (seperti pendidikan, latar belakang pekerjaan, serta demografi) maupun faktor makroekonomi yang mencerminkan kondisi bisnis dan ekonomi memiliki pengaruh signifikan terhadap tingkat *underemployment*. Selain itu, penelitian tersebut juga menunjukkan bahwa pekerja yang berada dalam kondisi *underemployed* cenderung memiliki tingkat kesejahteraan yang lebih rendah, bahkan lebih berisiko masuk dalam kategori rumah tangga miskin. Temuan ini menegaskan bahwa *underemployment* tidak hanya berdampak pada pasar tenaga kerja, tetapi juga pada dimensi kesejahteraan sosial.

Di tingkat regional, studi oleh Salsabilla dan Aji (2025) menunjukkan bahwa PDRB memiliki pengaruh positif terhadap peningkatan jumlah tenaga kerja yang terserap. Namun, penelitian yang sama bersama dengan temuan dari Rahmi dan Riyanto (2022) menyoroti kompleksitas pengaruh Upah Minimum Provinsi (UMP). Kenaikan UMP yang pada dasarnya ditujukan untuk melindungi kesejahteraan pekerja justru dapat menimbulkan efek samping berupa efisiensi perusahaan, seperti pengurangan jam kerja atau pergeseran tenaga kerja ke sektor informal. Kondisi ini memperkuat asumsi bahwa kebijakan upah yang tidak diimbangi dengan peningkatan produktivitas dan keterampilan dapat memperburuk fenomena *underemployment* di daerah dengan basis ekonomi yang lemah.



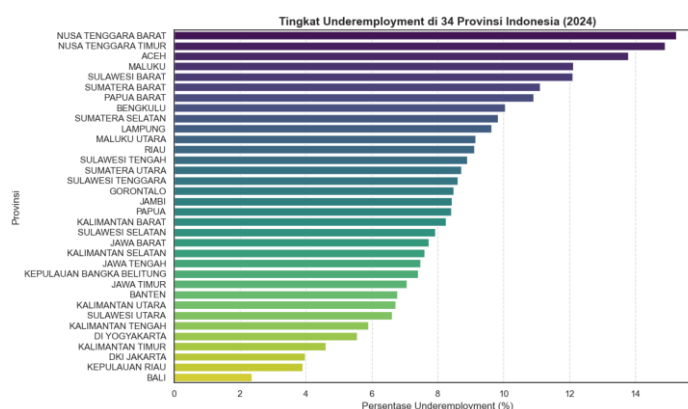
Gambar 2. Hubungan Indeks Daya Saing Digital terhadap *Underemployment*

Sumber: Hasil olahan penulis berdasarkan data BPS (2024)

Gambar diatas memperlihatkan adanya hubungan negatif antara Indeks Daya Saing Digital dan tingkat *underemployment* antar provinsi. Provinsi dengan skor digital *competitiveness* tinggi, seperti DKI Jakarta, DI Yogyakarta, dan Bali, cenderung memiliki tingkat *underemployment* yang lebih rendah. Sebaliknya, provinsi di kawasan timur Indonesia seperti Nusa Tenggara Timur, Nusa Tenggara Barat, dan Maluku masih menghadapi tingkat *underemployment* yang tinggi meskipun terjadi pertumbuhan ekonomi nasional.

Temuan ini menunjukkan bahwa peningkatan daya saing digital belum sepenuhnya merata di seluruh wilayah Indonesia. Akses infrastruktur digital dan kualitas SDM menjadi faktor penting yang memengaruhi kemampuan suatu daerah dalam menyerap tenaga kerja penuh. Oleh karena itu, pemerataan pembangunan digital dapat menjadi salah satu strategi kebijakan dalam menurunkan angka *underemployment*, terutama di wilayah yang tertinggal secara teknologi.

Kehadirna ekonomi digital memper rumit dinamika ketenagakerjaan di Indonesia. Sejumlah studi (Akmalya, 2023; Irawan et al., 2025) menyoroti peran ganda ekonomi digital, di satu sisi menciptakan peluang kerja baru melalui *e-commerce*, *startup*, dan ekonomi kreatif, namun di sisi lain mempopulerkan model *gig economy* yang fleksibel namun seringkali tidak menentu. Hal ini mengindikasikan bahwa Indeks Daya Saing Digital suatu provinsi tidak serta merta berkorelasi negatif dengan *underemployment*. Di sisi lain, faktor fundamental seperti Indeks Pembangunan Manusia (IPM), yang mencerminkan kualitas SDM, juga memiliki hubungan yang bervariasi tergantung pada kesesuaian keahlian dengan kebutuhan pasar kerja digital yang terus berubah.



Gambar 3. Sebaran *Underemployment* di 34 Provinsi Indonesia Tahun 2024
Sumber: BPS sakernas (2024)

Berdasarkan Gambar diatas, terlihat adanya variasi tingkat *underemployment* yang cukup lebar antar provinsi. Provinsi-provinsi di kawasan timur Indonesia, seperti Nusa Tenggara Timur dan Maluku, cenderung memiliki tingkat *underemployment* yang lebih tinggi dibandingkan provinsi-provinsi di Pulau Jawa. Hal ini dapat mencerminkan adanya ketimpangan dalam struktur ekonomi regional serta keterbatasan lapangan kerja formal. Sebaliknya, provinsi dengan tingkat digitalisasi tinggi seperti DKI Jakarta dan Jawa Barat menunjukkan tingkat *underemployment* yang relatif lebih rendah, meskipun tetap menghadapi tantangan dalam sektor informal.

Kondisi ini mengindikasikan bahwa penetrasi ekonomi digital dan kualitas infrastruktur teknologi informasi dapat berkontribusi dalam mengurangi tingkat pekerjaan tidak penuh, namun efeknya tidak bersifat merata di seluruh wilayah. Dengan demikian, variasi spasial ini memberikan dasar empiris bagi analisis lebih lanjut mengenai faktor-faktor yang memengaruhi *underemployment*. Pendekatan berbasis machine learning menjadi relevan karena mampu menangkap kompleksitas hubungan antara variabel-variabel ekonomi dan sosial yang berbeda antar wilayah.

Sebagian besar dari penelitian sebelumnya masih menggunakan pendekatan ekonometrika linier konvensional, seperti regresi OLS atau model efek tetap (Fixed Effect). Pendekatan tersebut cenderung mengasumsikan hubungan linear antar variabel dan sulit menangkap pola hubungan yang kompleks serta interaksi non-linier antar faktor ekonomi dan sosial. Padahal, dinamika *underemployment* di era digital kemungkinan besar bersifat multidimensional dan non-linier, melibatkan interaksi antara digitalisasi, pendapatan regional, pendidikan, dan kebijakan upah. Oleh karena itu, dibutuhkan pendekatan yang lebih adaptif dan mampu menangkap pola hubungan yang kompleks tersebut.

Untuk menjawab kesenjangan ini, penelitian ini menggunakan pendekatan machine learning berbasis data panel, khususnya model Random Forest, XGBoost, CatBoost, dan Gradient Boosting, untuk memprediksi tingkat *underemployment* di 34 provinsi Indonesia. Model ini mempertimbangkan variabel Indeks daya saing digital, PDRB per kapita, IPM, dan UMP sebagai determinan utama. Algoritma ensemble berbasis tree (seperti Random Forest dan varian Gradient Boosting)

dipilih karena kemampuannya yang terbukti dalam menangani data tabular, menangkap interaksi non-linier secara otomatis, serta ketahanannya terhadap outlier. Pendekatan *machine learning* memungkinkan model untuk belajar dari pola data historis, mengenali hubungan non-linier antar variabel, serta meningkatkan akurasi prediksi dibandingkan pendekatan ekonometrika tradisional.

B. Tujuan Penelitian

Penelitian ini bertujuan untuk menganalisis faktor-faktor yang memengaruhi tingkat *underemployment* di 34 provinsi Indonesia dengan mempertimbangkan peran ekonomi digital, PDRB per kapita, Indeks Pembangunan Manusia (IPM), dan Upah Minimum Provinsi (UMP). Melalui pendekatan *machine learning* berbasis data panel, penelitian ini berupaya mengidentifikasi pola hubungan non-linier dan kompleks antara variabel-variabel ekonomi dan sosial yang memengaruhi kualitas penyerapan tenaga kerja. Selain itu, penelitian ini juga bertujuan untuk mengembangkan model prediksi tingkat *underemployment* menggunakan beberapa algoritma *machine learning* seperti Random Forest, XGBoost, CatBoost, dan Gradient Boosting, serta mengevaluasi kinerja masing-masing model dalam menjelaskan variasi *underemployment* antar provinsi dan antar waktu. Hasil penelitian diharapkan dapat memberikan pemahaman empiris yang lebih mendalam mengenai paradoks penyerapan tenaga kerja di era digital serta menjadi dasar bagi penyusunan kebijakan berbasis data untuk meningkatkan kualitas pekerjaan di Indonesia.

C. Tinjauan Literatur

Fenomena *underemployment* di Indonesia telah menjadi topik penting dalam penelitian ekonomi ketenagakerjaan, karena menggambarkan kualitas pekerjaan yang tidak optimal di tengah pertumbuhan ekonomi yang relatif kuat. Pratomo (2015) menunjukkan bahwa *underemployment* dipengaruhi oleh karakteristik individu seperti pendidikan, usia, dan jenis pekerjaan, serta oleh faktor makroekonomi seperti pertumbuhan ekonomi dan dinamika pasar tenaga kerja. Penelitian tersebut juga menemukan bahwa pekerja dalam kondisi *underemployed* memiliki tingkat kesejahteraan lebih rendah dan cenderung masuk dalam kelompok rumah tangga miskin. Hal serupa diungkap oleh Maryati dkk (2021), yang meneliti hubungan antara pertumbuhan ekonomi dan penyerapan tenaga kerja di Sumatera

Barat. Meskipun pertumbuhan ekonomi cenderung meningkatkan kesempatan kerja, kualitas pekerjaan yang tercipta belum sepenuhnya memadai untuk mengurangi *underemployment*.

Secara teoretis, hubungan antara pertumbuhan ekonomi dan ketenagakerjaan dijelaskan melalui Hukum Okun (Okun's Law), yang menyatakan bahwa terdapat hubungan negatif antara pertumbuhan ekonomi dan pengangguran. Menurut Mankiw, peningkatan pertumbuhan ekonomi sebesar 2 persen dapat menurunkan tingkat pengangguran sekitar 1 persen (Maryati dkk, 2021). Dengan demikian, peningkatan pertumbuhan ekonomi diharapkan dapat memperluas kesempatan kerja dan menurunkan pengangguran maupun *underemployment*. Pertumbuhan ekonomi regional direpresentasikan melalui Produk Domestik Regional Bruto (PDRB), baik atas harga berlaku maupun konstan, yang masing-masing mencerminkan struktur ekonomi dan laju pertumbuhan riil suatu daerah (Tutupoho, 2019).

Hubungan antara PDRB dan penyerapan tenaga kerja juga telah dibahas dalam beberapa penelitian lain. Salsabilla dan Aji (2025) menemukan bahwa PDRB berpengaruh positif terhadap peningkatan penyerapan tenaga kerja di Jawa Timur, namun efeknya terhadap *underemployment* tidak selalu linier. Tangke (2023) menambahkan bahwa pertumbuhan ekonomi yang didominasi oleh sektor informal yang dapat menghambat penyerapan tenaga kerja secara penuh. Penelitian oleh Sipahutar dan Nasution (2023) juga mendukung pandangan tersebut dengan menunjukkan bahwa peningkatan PDRB per kapita dan IPM berkontribusi terhadap penurunan kemiskinan, tetapi masih belum cukup menekan tingkat pekerjaan tidak penuh di beberapa wilayah.

Selain faktor pertumbuhan ekonomi, kebijakan upah juga berpengaruh terhadap dinamika ketenagakerjaan. Rahmi dan Riyanto (2022) menunjukkan bahwa kenaikan Upah Minimum Provinsi (UMP) dapat meningkatkan produktivitas tenaga kerja di industri manufaktur, tetapi juga menimbulkan efek efisiensi berupa pengurangan jam kerja atau pergeseran tenaga kerja ke sektor informal. Hasil serupa ditemukan oleh Auliya dan Agusalm (2022) serta Wulandari dan Woyanti (2023), yang menyoroti bahwa upah memiliki hubungan yang kompleks dengan tingkat pengangguran, terutama di kalangan tenaga kerja terdidik,

di mana peningkatan upah dapat mendorong pergeseran ke sektor informal atau pekerjaan paruh waktu. Dengan demikian, kebijakan upah yang tidak diimbangi dengan peningkatan produktivitas justru dapat memperburuk fenomena *underemployment*.

Dalam konteks transformasi ekonomi digital, sejumlah penelitian menunjukkan dampak ganda dari perkembangan teknologi terhadap ketenagakerjaan. Akmalia (2023) menyoroti bahwa ekonomi digital berperan penting dalam mendorong pertumbuhan ekonomi nasional melalui *startup*, *e-commerce*, dan sektor kreatif, namun pada saat yang sama ekonomi digital juga memperluas model kerja fleksibel seperti *gig economy* yang tidak stabil secara pendapatan maupun jam kerja. Irawan dkk (2025) menemukan bahwa ekonomi digital di Indonesia pada periode 2010–2023 berkontribusi terhadap penciptaan lapangan kerja baru, tetapi mayoritas bersifat sementara dan berbasis proyek. Rosadi dan Tomy (2025) serta Sari dkk (2025) juga menyatakan bahwa digitalisasi ekonomi membuka peluang bagi pekerja muda, tetapi berpotensi meningkatkan ketimpangan pekerjaan jika tidak diimbangi dengan peningkatan keterampilan digital.

Sebagian besar studi tersebut masih menggunakan pendekatan ekonometrika tradisional seperti regresi linier berganda (*Ordinary Least Squares/OLS*) atau model efek tetap (*Fixed Effect Model*). Pendekatan tersebut mengasumsikan hubungan antar variabel bersifat linier dan tidak mampu menangkap dinamika non-linier serta interaksi kompleks antara digitalisasi, pembangunan manusia, dan kebijakan ekonomi. Padahal, perubahan struktur ekonomi akibat digitalisasi kemungkinan besar menimbulkan pola hubungan yang tidak konstan antar provinsi dan antar waktu.

Untuk mengatasi keterbatasan tersebut, penelitian ini mengadopsi pendekatan *machine learning* berbasis data panel, menggunakan algoritma seperti Random Forest, XGBoost, CatBoost, dan Gradient Boosting. Pendekatan ini memungkinkan model mengenali pola hubungan non-linier antar variabel serta meningkatkan akurasi prediksi tingkat *underemployment* di 34 provinsi Indonesia. Dengan demikian, penelitian ini berupaya memberikan kontribusi baru baik secara metodologis, dengan penerapan *machine learning* dalam konteks ekonomi tenaga

kerja maupun secara empiris, melalui identifikasi faktor-faktor yang paling berpengaruh terhadap *underemployment* di era ekonomi digital.

Dari hasil tinjauan literatur, dapat diimplikasikan bahwa pertumbuhan ekonomi, tingkat digitalisasi, kebijakan upah, dan kualitas sumber daya manusia memiliki pengaruh terhadap variasi tingkat *underemployment* di Indonesia, namun arah dan kekuatannya belum tentu linier. Oleh karena itu, penelitian ini berhipotesis secara implisit bahwa hubungan antara variabel-variabel tersebut bersifat kompleks dan dapat dijelaskan lebih baik melalui pendekatan *machine learning* berbasis data panel yang mampu menangkap pola non-linier dan interaksi multidimensional antar faktor ekonomi dan sosial.

Tabel 1. Penelitian Terdahulu

Peneliti	Tahun	Metode	Variabel Utama	Temuan Utama	Gap Penelitian
Pratomo	2015	Regresi Multinomial Logit	Karakteristik individu, variabel makro	Faktor individu dan makro berpengaruh terhadap <i>underemployment</i>	Data <i>cross-sectional</i> ; belum mempertimbangkan digitalisasi dan dinamika waktu
Maryati dkk	2021	Regresi Linier	PDRB, UMP, IPM	Pertumbuhan ekonomi meningkatkan penyerapan tenaga kerja	Fokus pada kuantitas tenaga kerja; kualitas pekerjaan belum dianalisis
Auliya & Agusalim	2022	Regresi Panel (Fixed Effect)	COVID-19, Upah Riil, UMP Riil, PDRB per kapita, IPM, Keterbukaan Ekonomi	Upah riil dan UMP berpengaruh terhadap pengangguran terdidik, dengan efek berbeda sebelum dan sesudah pandemi	Analisis hanya pada pengangguran terdidik; belum membahas <i>underemployment</i> ; hubungan non-linier belum diuji
Rahmi & Riyanto	2022	Regresi Panel (FE)	UMP, produktivitas	UMP meningkatkan produktivitas tapi	Hubungan non-linier antarvariabel belum dieksplorasi

				mengurangi jam kerja	
Akmalya	2023	Deskriptif Kualitatif	Ekonomi digital, e-commerce	Ekonomi digital mendorong pertumbuhan ekonomi	Tidak menganalisis dampaknya terhadap <i>underemployment</i>
Irawan	2025	Studi literatur	Digitalisasi, ketenagakerjaan	Digitalisasi menciptakan pekerjaan baru tapi fleksibel	Belum diuji secara kuantitatif dan antar provinsi
Anggara & Auwalin	2025	<i>Propensity Score Matching (PSM)</i> + Analisis Sensitivitas Rosenbaum	Partisipasi Kartu Prakerja, status pekerjaan, jam kerja	Program Kartu Prakerja meningkatkan peluang kerja peserta muda, tapi belum efektif mengatasi <i>underemployment</i>	Menggunakan data <i>cross-sectional</i> ; belum menangkap efek jangka panjang; panel data diperlukan untuk melihat dinamika pekerjaan

D. Metode Penelitian

Penelitian ini menggunakan pendekatan kuantitatif eksplanatori dengan tujuan untuk menganalisis hubungan antara variabel-variabel ekonomi dan sosial yang memengaruhi tingkat *underemployment* di Indonesia. Pendekatan ini dipilih karena mampu memberikan bukti empiris atas hubungan antar variabel secara terukur sekaligus memungkinkan penerapan metode prediksi yang bersifat non-linier. Penelitian ini mengombinasikan karakteristik data lintas waktu (*time series*) dan antar wilayah (*cross section*) dalam bentuk data panel yang mencakup 34 provinsi di Indonesia selama periode analisis beberapa tahun terakhir, sehingga dapat menangkap dinamika spasial dan temporal dari *underemployment* di tingkat regional.

Data yang digunakan merupakan data sekunder yang bersumber dari publikasi resmi Badan Pusat Statistik (BPS), Kementerian Ketenagakerjaan, serta Kementerian Komunikasi dan Informatika (Kominfo). Variabel utama meliputi tingkat *underemployment* sebagai variabel dependen, sedangkan variabel independen terdiri atas Indeks Daya Saing Digital, Produk Domestik Regional Bruto (PDRB) per kapita, Indeks Pembangunan Manusia (IPM), dan Upah

Minimum Provinsi (UMP). Selain itu, dilakukan standarisasi terhadap variabel numerik agar setiap indikator memiliki skala yang sebanding dalam proses pelatihan model *machine learning*. Secara operasional, setiap variabel penelitian didefinisikan sebagai berikut:

Tabel 2. Variabel yang digunakan

Variabel	Jenis	Definisi Operasional	Satuan / Skala	Sumber Data
<i>Underemployment</i>	Dependen	Persentase penduduk bekerja <35 jam/minggu dan masih menginginkan jam kerja tambahan	Persen (%)	BPS (Sakernas)
Indeks Daya Saing Digital	Independen	Ukuran daya saing digital suatu provinsi berdasarkan indikator infrastruktur, SDM, dan adopsi teknologi	Indeks (0–100)	Kementerian Kominfo
PDRB per Kapita (ADHK)	Independen	Nilai tambah bruto riil per kapita pada harga konstan, mencerminkan tingkat kesejahteraan dan produktivitas	Rupiah (konstan 2010)	BPS
Indeks Pembangunan Manusia (IPM)	Independen	Indeks komposit yang mengukur kualitas hidup berdasarkan pendidikan, kesehatan, dan standar hidup layak	Indeks (0–100)	BPS
Upah Minimum Provinsi (UMP)	Independen	Besaran upah minimum yang ditetapkan oleh pemerintah provinsi setiap tahun	Rupiah	Kementerian Ketenagakerjaan
Periode / Tahun	Kontrol	Tahun pengamatan untuk setiap provinsi dalam dataset panel	Tahun (t)	BPS & Kominfo

Tabel tersebut menunjukkan bahwa variabel-variabel yang digunakan mencerminkan kombinasi antara faktor ekonomi, sosial, dan digitalisasi yang dianggap memengaruhi kualitas penyerapan tenaga kerja. Dengan komposisi variabel seperti ini, penelitian diharapkan mampu menggambarkan hubungan

multidimensi yang terjadi antar provinsi di Indonesia, baik secara spasial maupun temporal.

Model riset yang digunakan dalam penelitian ini bertujuan untuk memprediksi variasi *underemployment* berdasarkan kombinasi faktor ekonomi, sosial, dan digital menggunakan pendekatan *machine learning* berbasis data panel. Empat model pembelajaran yang digunakan adalah *Random Forest*, *XGBoost*, *CatBoost*, dan *Gradient Boosting Regressor*. Pemilihan keempat model ini didasarkan pada kemampuannya menangani hubungan non-linier, interaksi kompleks antar variabel, serta keunggulan dalam mengatasi *overfitting* melalui teknik regularisasi dan *ensemble learning*.

Dalam tahap analisis, data dibagi menjadi set pelatihan (*train*) dan set pengujian (*test*) dengan rasio 80:20 untuk memastikan generalisasi model. Setiap model kemudian dievaluasi menggunakan metrik R^2 (koefisien determinasi), Adjusted R^2 , Mean Absolute Error (MAE), Mean Squared Error (MSE), Root Mean Squared Error (RMSE), Mean Absolute Percentage Error (MAPE), serta Explained Variance Score. Hasil evaluasi tiap algoritma dibandingkan untuk menentukan model dengan tingkat akurasi dan kemampuan prediksi terbaik terhadap variasi *underemployment* antar provinsi di Indonesia.

Pendekatan ini diharapkan dapat menangkap hubungan yang lebih kompleks dibandingkan metode *ekonometrika* tradisional seperti OLS atau Fixed Effect Model. Dengan demikian, penelitian ini tidak hanya memberikan gambaran empiris mengenai faktor-faktor yang memengaruhi *underemployment*, tetapi juga menawarkan model prediksi berbasis data yang dapat digunakan untuk perumusan kebijakan ketenagakerjaan di era ekonomi digital.

E. Hasil dan Analisis Pembahasan

Dalam penelitian ini data yang digunakan mencakup 34 provinsi di Indonesia dengan rentang waktu pengamatan selama 5 tahun (2020-2024). Variabel yang digunakan ialah Indeks Daya Saing Digital, Produk Domestik Regional Bruto (PDRB) per kapita, Indeks Pembangunan Manusia (IPM), Upah Minimum Provinsi (UMP), serta tingkat *underemployment* sebagai variabel dependen. Seluruh variabel dinormalisasi menggunakan metode Robust Scaler untuk mengatasi potensi keberadaan outlier dan menjaga kestabilan model. PDRB per kapita digunakan

sebagai representasi pertumbuhan ekonomi regional, sedangkan IPM mencerminkan kualitas sumber daya manusia. UMP menggambarkan kebijakan upah di tingkat provinsi, dan Indeks Daya Saing Digital menjadi indikator tingkat adopsi dan kesiapan digital daerah.

Tabel 3. Statistika Deskriptif

Statistik	Underemployment	Indeks Daya Saing Digital	PDRB per kapita	IPM	UMP (Rp)	Tahun
N (count)	170	170	170	170	170	170
Mean	8.59	37.85	74,914.61	72.23	2,806,050	2022
Std. Dev.	2.92	10.65	57,774.73	3.71	591,008	1.42
Min	2.35	17.70	20,057	60.44	1,704,608	2020
25% (Q1)	6.41	30.93	43,763.25	70.32	2,448,299	2021
Median (Q2)	8.38	35.90	56,844	72.17	2,805,248	2022
75% (Q3)	10.43	42.60	73,843.50	73.48	3,166,314	2023
Max	16.83	79.70	344,350	83.08	5,067,381	2024

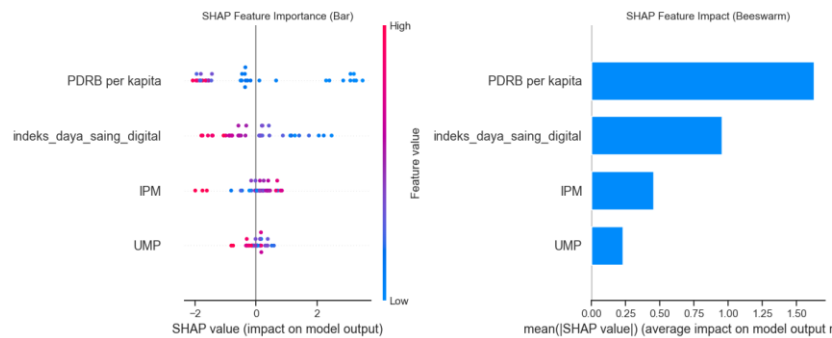
Dari tabel di atas terlihat bahwa rata-rata tingkat *underemployment* di Indonesia selama periode observasi adalah 8,59% dengan variasi antar provinsi yang cukup tinggi. Hal ini menunjukkan bahwa fenomena pekerjaan tidak penuh masih cukup signifikan di sebagian besar daerah. Sementara itu, rata-rata Indeks Daya Saing Digital sebesar 37,85 menunjukkan bahwa sebagian besar provinsi berada pada tahap awal digitalisasi. Nilai IPM rata-rata sebesar 72,23 menggambarkan kualitas sumber daya manusia yang relatif baik, sedangkan rata-rata UMP sebesar Rp2,8 juta menunjukkan variasi kebijakan upah antar provinsi. Nilai PDRB per kapita memiliki variasi paling besar, dari Rp20.000,00 hingga Rp344.000,00 , yang menunjukkan ketimpangan ekonomi antar wilayah masih cukup lebar.

Setelah data bersih dan siap digunakan, dilakukan analisis menggunakan empat metode *machine learning*, yaitu Random Forest, XGBoost, CatBoost, dan Gradient Boosting, dengan tujuan membandingkan kemampuan masing-masing model dalam memprediksi tingkat *underemployment*.

Tabel 4. Perbandingan Model

Model	R ² Train	R ² Test	Adjusted R ²	RMSE	MAE	MAPE (%)	Overfitting
Random Forest	0.8253	0.7036	0.6627	1.659	1.262	18.82	Moderat
XGBoost	0.9142	0.7031	0.6622	1.661	1.323	18.91	Tinggi
CatBoost	0.9441	0.7702	0.7385	1.461	1.101	17.10	Moderat
Gradient Boosting	0.9588	0.6878	0.6447	1.703	1.329	19.14	Tinggi

Berdasarkan hasil pengujian di atas, model CatBoost menunjukkan performa terbaik dengan nilai R² Test sebesar 0.7702 dan Adjusted R² sebesar 0.7385. Artinya, sekitar 74% variasi tingkat *underemployment* antar provinsi dapat dijelaskan oleh variabel ekonomi dan sosial yang digunakan. Nilai Mean Absolute Percentage Error (MAPE) sebesar 17,10% menunjukkan tingkat kesalahan prediksi yang tergolong rendah, sehingga model memiliki tingkat akurasi yang baik.



Gambar 4. Analisis pengaruh variabel terhadap tingkat *underemployment*

Setelah mengetahui bahwa model CatBoost memberikan hasil terbaik dalam memprediksi tingkat *underemployment*, langkah selanjutnya adalah melakukan analisis kontribusi variabel independen terhadap variabel dependen menggunakan metode SHAP (*SHapley Additive exPlanations*). Berdasarkan gambar di atas, terlihat bahwa variabel PDRB per kapita memiliki pengaruh paling besar terhadap *underemployment*, yang berarti bahwa perubahan pada tingkat PDRB per kapita memberikan dampak signifikan terhadap variasi *underemployment* antar provinsi. Semakin tinggi nilai PDRB per kapita suatu daerah, umumnya tingkat *underemployment* cenderung menurun karena peningkatan aktivitas ekonomi mendorong penciptaan pekerjaan yang lebih produktif.

Selanjutnya, untuk variabel Indeks Daya Saing Digital dan Indeks Pembangunan Manusia (IPM) juga menunjukkan pengaruh yang cukup kuat. Hal ini menandakan bahwa peningkatan kemampuan digital suatu wilayah serta kualitas sumber daya manusia turut berkontribusi dalam menekan *underemployment*. Daerah dengan tingkat digitalisasi yang tinggi cenderung memiliki lebih banyak peluang kerja fleksibel dan berbasis teknologi, sedangkan IPM yang tinggi mengindikasikan kemampuan adaptasi tenaga kerja terhadap perubahan pasar kerja yang lebih baik.

Sementara itu, variabel Upah Minimum Provinsi (UMP) memiliki pengaruh paling kecil dibandingkan variabel lainnya. Meskipun UMP tetap berperan dalam menjaga kesejahteraan tenaga kerja, pengaruhnya terhadap *underemployment* relatif terbatas. Hal ini bisa disebabkan oleh adanya ketidaksesuaian antara tingkat upah dengan produktivitas tenaga kerja, atau karena sebagian besar pekerja berada di sektor informal yang tidak sepenuhnya terpengaruh oleh kebijakan upah minimum.

Secara keseluruhan, hasil analisis SHAP ini memperkuat temuan bahwa faktor ekonomi makro seperti PDRB per kapita serta faktor struktural seperti digitalisasi dan kualitas sumber daya manusia memiliki peran penting dalam menekan tingkat *underemployment* di Indonesia, sementara kebijakan upah perlu disertai dengan peningkatan produktivitas agar dapat memberikan efek yang lebih optimal. Secara keseluruhan, hasil ini sejalan dengan temuan Rahmi dan Riyanto, (2022) serta Auliya dan Agusalm (2022), yang menunjukkan bahwa kebijakan upah dan faktor makroekonomi memiliki hubungan kompleks terhadap kualitas pekerjaan.

F. Penutup

Penelitian ini berhasil menganalisis dan memprediksi faktor-faktor yang memengaruhi tingkat *underemployment* di 34 provinsi Indonesia dengan menggunakan pendekatan *machine learning* untuk menangkap hubungan non-linier yang kompleks. Hasil perbandingan menunjukkan bahwa model CatBoost memiliki kinerja prediksi terbaik (R^2 Test 0.7702 dan MAPE 17.10%), yang mampu menjelaskan sekitar 74% variasi *underemployment* antar provinsi. Analisis kontribusi variabel menggunakan SHAP mengungkap temuan kunci bahwa PDRB

per kapita adalah faktor prediktif paling dominan, menegaskan bahwa pertumbuhan ekonomi regional tetap menjadi motor utama dalam penyerapan tenaga kerja. Selain itu, Indeks Daya Saing Digital (IED) dan Indeks Pembangunan Manusia (IPM) juga menunjukkan pengaruh yang kuat. Temuan ini secara langsung menjawab paradoks penelitian: pertumbuhan ekonomi (PDRB) saja tidak cukup menekan *underemployment* jika tidak diimbangi dengan peningkatan kualitas SDM (IPM) dan kesiapan infrastruktur serta adopsi teknologi (IED). Sebaliknya, Upah Minimum Provinsi (UMP) memiliki daya prediktif paling kecil, mengindikasikan bahwa kebijakan upah minimum memiliki keterbatasan dalam mengatasi masalah *underemployment* yang bersifat struktural, kemungkinan besar karena banyaknya pekerja di sektor informal.

Temuan ini memberikan implikasi kebijakan yang penting bagi pemerintah pusat maupun daerah. Untuk mengatasi paradoks *underemployment* secara efektif, fokus kebijakan tidak bisa lagi bersifat parsial. Pemerintah daerah harus fokus mendorong pertumbuhan ekonomi yang berkualitas (peningkatan PDRB) yang mampu menciptakan lapangan kerja formal dan produktif, bukan sekadar memperluas *gig economy* atau sektor informal. Sejalan dengan itu, investasi pada modal manusia (IPM) dan infrastruktur digital (IED) harus menjadi prioritas. Peningkatan keterampilan digital dan kemampuan adaptasi angkatan kerja adalah kunci agar mereka dapat terserap ke dalam pekerjaan yang lebih baik dan penuh waktu. Terakhir, kebijakan UMP perlu ditinjau ulang agar efektivitasnya sejalan dengan peningkatan produktivitas, sehingga tidak justru mendorong perusahaan melakukan efisiensi dengan mengurangi jam kerja.

Penelitian ini memiliki beberapa keterbatasan. Pertama, penggunaan data agregat di level provinsi mungkin menyembunyikan variasi yang signifikan di tingkat kabupaten/kota. Kedua, model *machine learning* sangat baik dalam prediksi dan identifikasi pentingnya fitur, namun tidak dapat membuktikan hubungan kausalitas (sebab-akibat) secara formal. Untuk penelitian selanjutnya, disarankan menggunakan data panel mikro (data mentah Sakernas) untuk menganalisis bagaimana karakteristik individu (seperti tingkat pendidikan, usia, jenis kelamin, dan sektor pekerjaan) berinteraksi dengan faktor makro dan digitalisasi dalam menentukan status *underemployment* seseorang.