

Pacmann

Komparasi Performa Model Machine Learning untuk Memprediksi Nilai Ecological Footprint Per Kapita Suatu Negara

FINAL PROJECT: INTRODUCTION TO MACHINE LEARNING

Nuzula Elfa Rahma

I. Pendahuluan

Jika kita menengok hingga beberapa dekade ke belakang, diskursus tentang kelestarian lingkungan cenderung dibingkai dengan keresahan tentang ledakan populasi penduduk dunia. Kecenderungan ini bisa kita temukan entah itu dalam perbincangan santai, diskusi akademik, artikel ilmiah, produk budaya, bahkan dalam buku pelajaran. Teori Malthus tentang tren pertumbuhan penduduk, dimana pertumbuhan populasi akan selalu cenderung melebihi persediaan makanan dan bahwa kemajuan umat manusia tidak mungkin terjadi tanpa pembatasan yang ketat pada aspek reproduksi, dipandang sebagai model yang mumpuni dalam hal memprediksi keruntuhan suatu institusi kemasyarakatan di masa depan. Dalam konteks keberlanjutan lingkungan, terma “persediaan makanan” melebar lingkupnya hingga mencakup segala sumberdaya yang diekstraksi dari lingkungan baik itu bahan makanan, bahan bakar, ataupun bahan mentah (Hasnain et al., 2005). Dengan menitikberatkan pada variabel jumlah dan pertumbuhan penduduk, tentu saja negara-negara berkembang lah yang cenderung dianggap sebagai kontributor utama penyebab degradasi lingkungan (Trainer, 1990).

Seiring dengan semakin mencuatnya isu perubahan iklim akibat emisi gas rumah kaca dari berbagai aktivitas manusia, konsep Malthus ini mulai mendapat banyak pertentangan dan sanggahan. Diskursus mengenai siapa yang patut disalahkan dalam isu kerusakan lingkungan terutama perubahan lingkungan mulai bergeser dan tertuju pada peran negara-negara maju (Post et al., 2019). Sejalan dengan semakin menonjolnya konsep jejak ekologis (*ecological footprint*) dan jejak karbon (*carbon footprint*), dimana kontribusi manusia dalam membebani kapasitas lingkungan kini sudah dapat dikuantifikasi, sehingga tidak lagi bersifat retorik. Adapun *ecological footprint* merupakan satuan (Global Hectare) yang mencerminkan banyaknya sumberdaya yang dikonsumsi oleh seseorang atau kumpulan masyarakat, dalam melangsungkan hidupnya baik itu dalam bentuk aktivitas pemenuhan kebutuhan pangan, sandang, listrik, transportasi, hiburan, komunikasi, hingga limbah yang dihasilkan. Sehingga upaya mitigasi yang lebih terukur dengan menyasar kepada fenomena, aktivitas dan gaya hidup yang benar-benar menjadi kontributor dampak lingkungan sudah dapat dilakukan. Salah satu upaya yang dapat dilakukan adalah dengan memanfaatkan model sebagai instrumen untuk melakukan simulasi ataupun prediksi magnitude dari aktivitas manusia di suatu wilayah terhadap lingkungannya berdasarkan parameter yang relevan.

Dengan latar belakang tersebut, maka objektif dari *final project* ini adalah membangun model berbasis *machine learning* yang dapat memprediksi nilai *ecological footprint* penduduk suatu negara, dengan menggunakan input demografi (misalnya seperti HDI dan harapan hidup) dan fitur kewilayahan suatu negara. Pendekatan yang diambil adalah dengan membandingkan antara model regresi linear dan regresi polinomial.

II. Studi Terkait

Dalam sebuah studi, (Wackernagel et al., 2017) menunjukkan bahwa kontradiktif dengan teori Malthus, negara-negara berpenduduk rendah namun dengan tingkat HDI (Human Development Index) tertinggi sedunia ternyata justru memiliki *ecological footprint* (EF) paling tinggi. Hasil studi ini juga mendapat perhatian dari Uni Eropa (European Environment Agency, 2019).

Dalam studi lainnya (Dietz et al., 2007) juga menyatakan bahwa pencetus utama dalam terjadinya kerusakan lingkungan bukan hanya jumlah populasi namun juga gaya hidup suatu populasi. Namun Dietz juga mengemukakan bahwa tingkat pendidikan dan harapan hidup tidak terlalu signifikan berkontribusi terhadap kerusakan lingkungan. Hal ini penting dicatat, karena tingkat pendidikan dan harapan hidup adalah variabel yang diperhitungkan dalam kalkulasi HDI. Sehingga bisa diasumsikan, bahwa tidak semua bagian dari HDI berkorelasi dengan meningkatnya EF suatu negara. Jika ditilik dari parameter ekonomi lain yang beririsan dengan HDI seperti Economic Complexity Index dimana negara-negara maju cenderung memiliki skor tinggi, studi dari Neagu (2020) juga menemukan bahwa ada hubungan asosiasi positif antara indeks tersebut dengan besaran EF dalam aspek produksi suatu negara. Secara umum berbagai studi cenderung sepakat bahwa populasi memang bukan satu-satunya pendorong utama kerusakan lingkungan, namun ada berbagai faktor lain yang juga ikut berperan.

III. Dataset dan Fitur

3.1. Dimensi dan struktur

Dataset yang digunakan dalam proyek berikut merupakan dataset time series tahunan dengan rentang waktu dari tahun 2000 sampai dengan 2014 yang mencakup data profil negara-negara seluruh dunia. Total entry yang termaktub dalam dataset tersebut adalah 2156 poin data. Dataset merupakan data repository Github yang diolah dengan menggabungkan data dari World Bank dan data dari Global Footprint Network (Shropshire, 2019). Profil tersebut terdiri dari beberapa variabel yaitu:

- *Country*:
Mencakup 146 negara di seluruh dunia.
- *Continent*:
Mencakup 6 wilayah benua yaitu Asia, Europe, Africa, South America, Oceania, dan North America.
- *HDI (Human Development Index)*
Parameter kependudukan yang merangkum tingkat kesejahteraan masyarakat suatu wilayah dari aspek kesehatan, pendidikan, dan taraf hidup masyarakat, dengan skala 0 – 1.
- *Life Expectancy*
Rata-rata harapan hidup penduduk suatu negara.
- *Population*
Jumlah penduduk suatu negara pada tahun ke-n.
- *Ecological Footprint per Capita (EFConsPerCap)*
Rata-rata besaran porsi kapasitas lingkungan per tahun yang dihabiskan per penduduk suatu negara dalam melangsungkan kehidupannya, mencakup semua kebutuhan baik primer, sekunder, ataupun tersier (Global Hectare per capita).
- *Total Ecological Footprint dalam Global Hectare (EFConsTotGHA)*.
Total besaran porsi kapasitas lingkungan yang dihabiskan suatu negara (Total GHA).
- *Biocapacity per Capita (BiocapPerCap)*
Kapasitas lingkungan suatu negara dalam penyediaan sumberdaya dalam pemenuhan gaya hidup per penduduk (GHA per capita).

- Total Biocapacity dalam Global Hectare (BiocapTotGHA)
Total kapasitas lingkungan suatu negara dalam penyediaan sumberdaya dalam pemenuhan gaya hidup per penduduk (GHA per capita).

Unnamed: 0	country	year	hdi	life_expectancy	population	continent	EFConsPerCap	EFConsTotGHA	BiocapPerCap	BiocapTotGHA
0	Afghanistan	2000	0.340	55.482	20093756.0	Asia	0.648085	12768498.74	0.509767	1.004337e+07
1	Afghanistan	2001	0.341	56.044	20966463.0	Asia	0.605820	12438205.89	0.500063	1.026689e+07
2	Afghanistan	2002	0.373	56.637	21979923.0	Asia	0.704061	15128227.46	0.570591	1.226035e+07
3	Afghanistan	2003	0.381	57.250	23064851.0	Asia	0.708856	15954511.65	0.591033	1.330261e+07
4	Afghanistan	2004	0.396	57.875	24118979.0	Asia	0.619414	14556109.86	0.502825	1.181630e+07
...
2152	Zimbabwe	2010	0.452	52.975	14086317.0	Africa	1.241330	17346227.76	0.619689	8.659478e+06
2153	Zimbabwe	2011	0.464	54.800	14386649.0	Africa	1.306024	18618155.97	0.580174	8.270734e+06
2154	Zimbabwe	2012	0.488	56.516	14710826.0	Africa	1.173113	17086982.45	0.587819	8.561875e+06
2155	Zimbabwe	2013	0.498	58.053	15054506.0	Africa	1.140062	16984754.24	0.558680	8.323270e+06
2156	Zimbabwe	2014	0.507	59.360	15411675.0	Africa	1.088133	16589564.28	0.540845	8.245675e+06

2157 rows × 11 columns

Gambar 1. Cuplikan dataset yang digunakan

Pada proyek ini, target yang difokuskan adalah pada variabel Ecological Footprint per Capita saja.

3.2. Training, validasi dan pengujian

Dataset dibagi dua menjadi data untuk keperluan training dan pengujian. Proporsi pembagian data adalah 80% untuk *training data*, dan 20% untuk *test data*. Adapun untuk keperluan validasi, data yang digunakan diambil dari data training, yaitu dengan proporsi 20% dari data training.

3.3. Preprocessing

Proses preprocessing yang dilakukan pada dataset di atas meliputi transformasi data kategorikal dan numerikal, dan standarisasi/normalisasi data input. Transformasi data kategorikal adalah pada kolom 'continent' dengan menggunakan metode transformasi *one hot encoding*. Transformasi ini akan menambah dimensi fitur dataset sesuai dengan jumlah *unique value* pada kolom kategorikal. Agar setiap fitur pada dataset memiliki distribusi yang sama, maka setelah *preprocessing*, dataset dinormalisasi terlebih dahulu. Dengan demikian, nilai-nilai pada dataset berada pada skala yang sama.

IV. Metode

4.1. Eksplorasi data

Eksplorasi data dilakukan untuk melihat hubungan antar berbagai variabel, baik antar input variabel maupun input-output variabel. Eksplorasi ini dilakukan dengan bantuan visualisasi data dalam bentuk plot regresi dan boxplot. Dari eksplorasi ini diharapkan banyaknya variabel yang relevan dapat dikerucutkan sebelum membangun model, sehingga mengurangi kompleksitas model yang tidak perlu.

4.2. Model Machine Learning

Dalam proyek ini, secara umum adalah melakukan eksperimentasi komparasi antara model regresi linear dan regresi polynomial. Selain itu juga dilakukan *feature engineering* dengan metode PCA (Principal Component Analysis).

a) Regresi linear

Regresi linear adalah salah satu model dasar dalam *machine learning* yang digunakan untuk membuat prediksi berdasarkan satu atau beberapa input variabel (fitur). Model ini mengasumsikan adanya hubungan linear (garis lurus) antara fitur dan target prediksi.

Regresi linear dapat dinotasikan sebagai berikut:

$$y = \beta_0 + \beta_1 x_1 + \beta_2 x_2 + \dots + \beta_n x_n$$

Di mana,

y	:	target yang ingin diprediksi
x_1, x_2, \dots, x_n	:	fitur yang digunakan sebagai input
$\beta_1, \beta_2, \dots, \beta_n$:	koefisien yang dicari saat melakukan proses <i>learning</i>
β_0	:	intercept (nilai default yang tetap tercapai meski tidak ada input)

b) Regresi polinomial

Berbeda dengan linear, regresi polinomial memperhitungkan interaksi antar fitur dalam memperhitungkan target. Dalam model ini, fitur yang ada juga dapat diberikan nilai pangkat tertentu, sehingga garis persamaannya cenderung berbentuk kurva.

Regresi polinomial dapat dinotasikan sebagai berikut:

$$y = \beta_0 + \beta_1 x_1 + \beta_2 x_2^2 + \beta_3 x_3^3 + \dots + \beta_n x_n^n$$

Di mana:

y	:	target yang ingin diprediksi
x_1, x_2, \dots, x_n	:	fitur yang digunakan sebagai input
$\beta_1, \beta_2, \dots, \beta_n$:	koefisien yang dicari saat melakukan proses <i>learning</i>
pangkat (2, 3, ..., n)	:	tingkat interaksi antar fitur
β_0	:	intercept (nilai default yang tetap tercapai meski tidak ada input)

Dalam pengaplikasiannya pada saat melakukan training, regresi polynomial dibangun dari beberapa fraksi model regresi linear dengan *slope* yang berbeda, sesuai dengan bentuk kurva polinomial model. Sehingga sebelum training, akan dibuat model *pipeline* yang fungsinya mengintegrasikan regresi linear ke dalam regresi polinomial. Dan model pipeline inilah yang nantinya akan difitting dengan *training data*.

c) Principal Component Analysis (PCA)

Transformasi data kategorikal memiliki konsekuensi yaitu membengkaknya jumlah fitur input, yang tentu saja akan menyebabkan semakin kompleksnya dimensi data yang harus diolah. Untuk itu metode *feature engineering* yakni PCA (Principal Component Analysis) diaplikasikan pada proyek ini. PCA adalah teknik mengurangi dimensi data dengan tetap mempertahankan sebanyak mungkin informasi penting yang dimiliki data tersebut. Secara prinsip, yang dilakukan dalam PCA adalah mengadopsi perspektif baru dengan tetap mengkomodir variansi terbesar pada data yang bersangkutan.

4.3. Training dan Prediksi

Proses Machine Learning memiliki dua tahap: training dan prediksi. Pada tahap pertama, model diberi data training untuk mempelajari bagaimana melakukan prediksi. Sebelum melakukan training, data training divalidasi menggunakan metode cross-validation dan 5 lipatan. Setiap lipatan memvalidasi 20% dari *data training* secara bergantian, menghasilkan metrik berupa nilai MSE. Rata-rata MSE diperiksa untuk menentukan apakah model memiliki risiko *overfitting* dengan membandingkannya dengan MSE baseline. Jumlah 5 lipatan dipilih karena efisien dan umum digunakan. Setelah itu, dilakukan *fitting* model lalu menjalankan model yang telah difit tersebut untuk melakukan prediksi pada training data

Tahap selanjutnya, prediksi dilakukan pada *data test* untuk mengevaluasi performa model yang telah dilatih dan memastikan bahwa model bekerja sesuai dengan harapan. Kali ini yang dilakukan adalah percobaan model yang telah difit serta dilatih pada training data untuk melakukan prediksi dimana inputnya adalah *test data*.

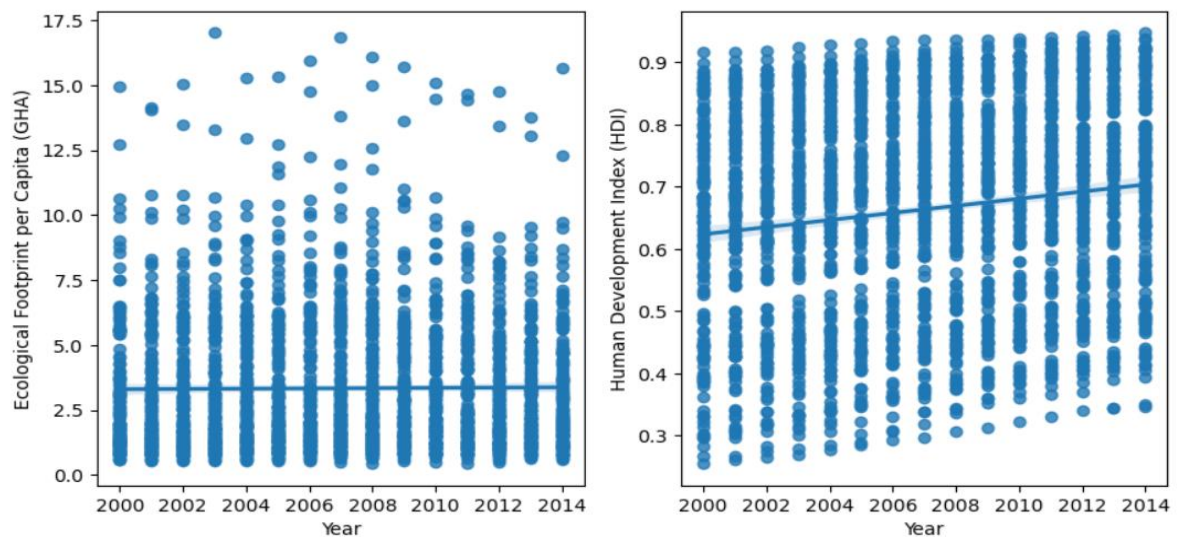
4.4. Evaluasi

Untuk menilai performa model, maka perlu dilakukan perbandingan performa terkait akurasi antara hasil prediksi training dan hasil prediksi tes. Metrik yang digunakan adalah MSE (Mean Squared Error). Secara prinsip, MSE adalah metrik yang digunakan untuk menilai kualitas model *machine learning* dengan mengukur rata-rata perbedaan antara nilai aktual dan nilai yang diprediksi oleh model. MSE dihitung dengan membagi jumlah kuadrat perbedaan antara target dan prediksi oleh jumlah data. Semakin kecil nilai MSE, semakin baik performa model dalam memprediksi target. MSE digunakan untuk menilai performa baik model training maupun model test. Dari nilai MSE ini maka dapat ditentukan berapa persen peningkatan akurasi prediksi jika dibandingkan baseline. Persentase ini akan diukur baik pada prediksi training maupun prediksi test, untuk kemudian dibandingkan.

V. Eksperimentasi, Hasil dan Pembahasan

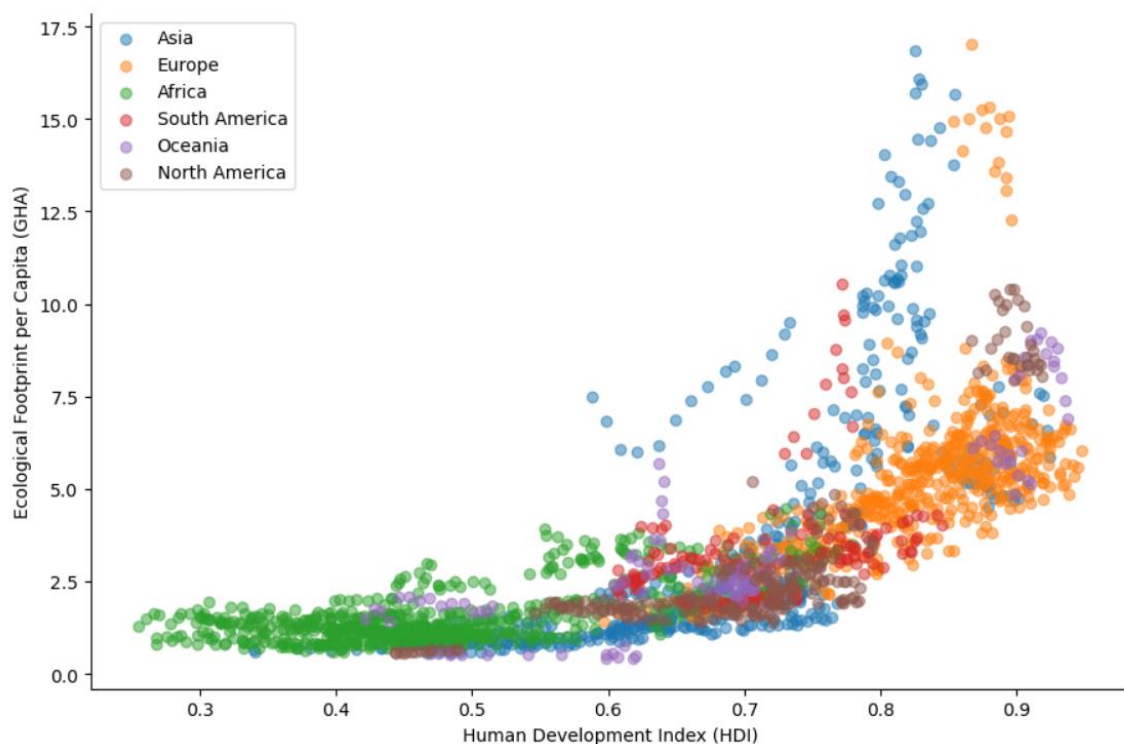
5.1. Eksplorasi data awal

Hal pertama yang menjadi pertimbangan dalam proyek ini adalah apakah penyajian dalam bentuk time series, diperlukan dalam membangun model ML dengan dataset ini. Apakah justru sebaiknya dikecualikan saja. Untuk menjawab hal tersebut, secara sederhana dapat dinilai dari korelasi antara fitur tahun (year) dengan HDI dan *ecological footprint per capita*.



Gambar 2. Korelasi antara (1) fitur periode tahun dan EF per kapita (2) fitur periode tahunan dan HDI

Dari plot regresi di atas (Gambar 2) dapat dilihat bahwa variabel waktu tahunan ternyata tidak menunjukkan korelasi terhadap EF per kapita. Sedangkan ketika diplotkan dengan fitur HDI, variabel periode tahunan menunjukkan korelasi positif namun tidak terlalu signifikan. Ini karena tren perubahan HDI cenderung membutuhkan jangka waktu beberapa dekade untuk dapat lebih terlihat polanya. Sebagai gambaran, satu studi berikut (Tiezzi et al., 2012) menggunakan data dengan jangka waktu dari tahun 1961 sampai dengan 2007. Atas dasar tersebut, maka variabel periode tahunan dikecualikan. Selain itu, fitur populasi dan rata-rata harapan hidup juga dikecualikan. Fitur populasi lebih memiliki hubungan dengan variabel Total EF suatu negara, ketimbang variabel EF per kapita. Selain itu dalam perhitungan HDI, angka harapan hidup merupakan salah satu variabel perhitungan (Mahajan, 2013).



Gambar 3. Plot regresi antara fitur HDI dan EF per kapita

Eksplorasi juga dilakukan untuk melihat apakah pendekatan model machine learning sudah berada di jalur yang sesuai. Dari Gambar 3 di atas, tampak bahwa ada hubungan korelasi positif antara nilai HDI dan EF per kapita. Dapat dilihat bahwa 70% data menunjukkan pola hubungan linear. Akan tetapi, bila dilihat lebih jauh pada data dengan HDI di kisaran di atas 0.8, ada data yang cenderung memiliki nilai EF per kapita dengan pola eksponensial. Selain itu dapat dilihat pula bahwa data Eropa cenderung berada di kisaran nilai EF yang tinggi, sementara Afrika cenderung di kisaran rendah. Sehingga diasumsikan bahwa model regresi polynomial dapat pula dijadikan pendekatan, karena adanya interaksi antar fitur. Atas dasar penilaian tersebut, maka dalam proyek ini, dilakukan eksperimentasi antar dua model yaitu regresi linear dan polynomial.

5.2. PCA

Dataset yang telah melalui *preprocessing*, memiliki dimensi 7 fitur. Untuk mengurangi dimensinya, dataset ini kemudian diolah dengan melakukan PCA. Dengan mempertahankan nilai variansi minimal 90%, dimensi data dapat dikurangi menjadi 5 fitur.

5.3. Evaluasi performa model

Seperti telah dijelaskan sebelumnya, evaluasi terkait performa model prediksi EF per kapita berdasarkan HDI dan region ini didasarkan pada perbandingan MSE antara hasil prediksi dan actual. Didapatkan MSE pada baseline training data adalah 7.27 dan pada baseline test data adalah 5.6. Performa model dikatakan baik apabila nilai MSE pada hasil prediksi, baik pada tahap training dan pengujian, lebih rendah daripada MSE baseline. Semakin jauh selisih nilainya maka semakin baik pula kualitas performanya.

Tabel 1. Performa model regresi linear dan regresi polynomial

No.	Pendekatan	Persentase penurunan MSE dari baseline	
		Training	Test
1.	Regresi polinomial (pangkat 2)	62%	-4957%
2.	Regresi linear (dengan PCA)	39.76%	52.07%
3.	Regresi linear (tanpa PCA)	53.47%	65.45%

Dapat dilihat pada Tabel 1 di atas, bahwa model regresi polynomial ternyata menunjukkan adanya *overfitting*. Performa pada saat training cukup baik, namun sangat buruk pada saat pengujian. Hasil *cross-validation score* mengindikasikan kecilnya peluang *overfitting*, dimana nilai error lebih kecil dari baseline. Namun ternyata hal itu tidak tercermin pada performa pengujian. Mengapa terjadi *overfitting*? Eksperimentasi dilakukan dengan berbagai nilai pangkat. Eksperimentasi ini menunjukkan semakin tinggi nilai pangkat yang diatribusikan, semakin buruk performa model. Hal yang sama juga terjadi dengan eksperimentasi mengurangi jumlah fitur. Semakin sedikit fitur, semakin baik performa pengujian. Sehingga dapat ditarik kesimpulan bahwa model regresi polinomial kurang tepat diaplikasikan karena terlalu kompleks.

Model regresi linear yang lebih sederhana ternyata menunjukkan performa yang jauh lebih baik. Dimana performa model pada saat training maupun pengujian menunjukkan hasil yang diharapkan. Namun yang menarik adalah, ketika dibandingkan performa antara dengan aplikasi PCA dan tanpa aplikasi PCA, nampak bahwa regresi linear tanpa PCA justru memiliki performa yang secara signifikan lebih baik. Hal ini mengindikasikan bahwasanya data dengan fitur yang sudah ada, pada dasarnya tidak memiliki masalah “*curse of dimensionality*”. Dalam artian, dimensinya tidak terlalu tinggi sehingga model masih mampu menangkap pola pada sebaran data. Tidak terlampau terdistraksi oleh *noise* pada data.

Sehingga dari beberapa pendekatan di atas, dapat dilihat bahwa model regresi linear tanpa perlakuan PCA *feature engineering* lah model yang paling akurat dalam melakukan prediksi EF per kapita berdasarkan fitur HDI dan benua.

5.4. Analisa Kelemahan pada Model

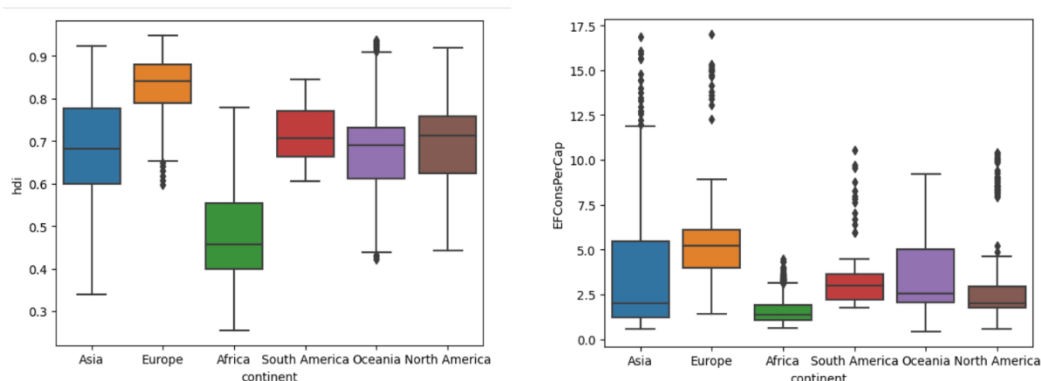
Meski regresi linear tanpa PCA menunjukkan performa yang paling baik, namun masih ada ruang untuk meningkatkan performa model tersebut. Beberapa aspek yang dapat dieksplorasi lebih jauh adalah sebagai berikut:

- *Periode waktu*

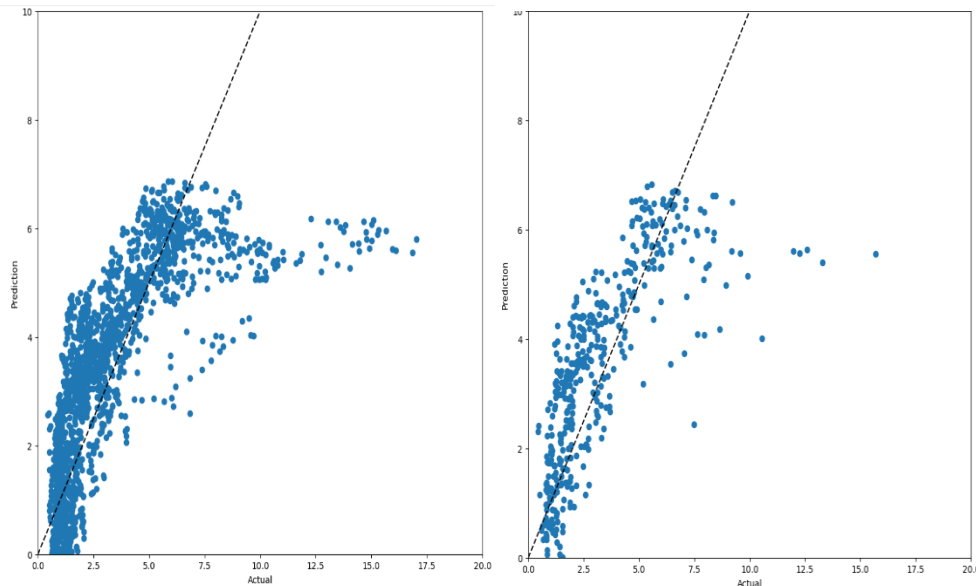
Dataset yang digunakan diolah dari beberapa dataset, dimana salah satu dataset yang berisi variabel EF per kapita merupakan *time series* dengan periode waktu dari tahun 1961 sampai tahun 2014, atau dalam jangka waktu 53 tahun. Pada jangka waktu ini, tren perubahan HDI dapat lebih dipetakan, ketimbang hanya 14 tahun.

- *Lingkup wilayah*

Pengklasifikasian wilayah ke dalam kategori benua adalah pengkerucutan dari kategori negara. Namun pengekerucutan ini dapat dikatakan terlalu sederhana dan tidak representatif. Dapat dilihat pada boxplot (Gambar 4) berikut, fitur benua pada kategori Eropa menunjukkan 70% data berada pada ranah HDI yang tinggi, tetapi tidak berdistribusi normal dan condong (*skewing*) sebelah kanan. Ini menunjukkan adanya outlier pada sebelah kiri. Hal serupa juga ditunjukkan pada hubungan antara fitur benua dan nilai EF per kapita, terlihat bahwa semua kategori memiliki kecenderungan condong ke sebelah kiri.



Gambar 4. Hubungan antara fitur benua dengan HDI dan EF per kapita



Gambar 5. Akurasi hasil prediksi pada data training (kiri) dan data pengujian (kanan) pada pendekatan regresi linear tanpa PCA

Dapat dilihat pula pada Gambar 5, bahwa kesalahan prediksi cenderung berada pada ranah 20% nilai tertinggi. Hal ini kemungkinan ada hubungannya dengan kumpulan *outlier* pada Gambar 4. Memang perlu dilakukan klasifikasi berbasis wilayah yang lebih representatif.

VI. Kesimpulan

Dari beberapa fitur yang terdapat di dataset, ada dua fitur yang dimasukkan sebagai input yaitu HDI dan “continent” atau benua dimana suatu negara berada. Setelah melalui tahapan *preprocessing* dan normalisasi, dua fitur ini berkembang dimensinya menjadi 7 fitur karena adanya data kategorikal. Ada beberapa pendekatan model *machine learning* yang dieksperimentasi, yaitu regresi polinomial dan regresi linear. Selain itu juga diterapkan PCA sebagai *feature engineering* dalam upaya mereduksi kompleksitas dimensi data. Dari beberapa pendekatan model machine learning untuk memprediksi EF per kapita berdasarkan fitur Human Development Index (HDI), dan benua (*continent*), berdasarkan performa seberapa besar model tersebut mampu menurunkan nilai MSE pada model baseline, didapatkan model yang terbaik adalah model regresi linear tanpa pengaplikasian *feature engineering* PCA (Principal Component Analysis). Performa pada sesi training dan pengujian tercatat di nilai 53.47% dan 65.45%. Jika dibandingkan dengan pendekatan lain, regresi polinomial menunjukkan adanya *overfitting*, dan aplikasi PCA justru menurunkan performa model.

Meski demikian, jika dilihat dari persentase tersebut, masih terdapat ruang yang cukup luas untuk dapat meningkatkan performa model tersebut. Konsep dalam pemutakhiran model prediksi EF per kapita adalah: 1) membangun dataset dengan fitur dan target yang sama namun dengan format time series dari tahun 1961 sampai dengan 2014, 2) memperkecil ruang lingkup wilayah pada skala regional (ex: Asia Tenggara, Asia Timur, Mediterania, dsb) sehingga data kategorikal “country” atau negara tidak memuat terlalu banyak *unique value* yang akan menambah kompleksitas dimensi data namun tetap representatif. Pendekatan

model yang akan digunakan adalah model yang memang lebih tepat untuk memprediksi data time series, misalnya model ARIMA. Selain itu, pengembangan model juga bisa diarahkan untuk memprediksi seberapa besar suatu negara telah menghabiskan alokasi biokapasitas lingkungan (baik secara per kapita maupun total biokapasitas). Dapat pula dikembangkan berbagai skenario tren perkembangan EF per kapita atau HDI suatu negara. Dan dari masing-masing skenario tersebut, kemudian diprediksi berapa lama waktu yang dibutuhkan suatu negara untuk melampaui biokapasitas negara tersebut.

VII. Referensi

- Dietz, T., Rosa, E. A., & York, R. (2007). Driving the Human Ecological Footprint. *Frontiers in Ecology and the Environment*, 5(1), 13–18. <http://www.jstor.org/stable/20440554>
- European Environment Agency. (2019). The European environment - state and outlook 2020. *European Environment Agency*.
- Hasnain, M., Azhar, U., Wasti, S., & Inam, Z. (2005). Interaction between Population and Environmental Degradation. *Pakistan Development Review*, 44. <https://doi.org/10.30541/v44i4Ipp.1135-1150>
- Mahajan, S. (2013). *Human Development Index — Measurements, changes and evolution*. 1–5. <https://doi.org/10.1109/NUiCONE.2013.6780196>
- Neagu, O. (2020). Economic Complexity and Ecological Footprint: Evidence from the Most Complex Economies in the World. *Sustainability*, 12(21). <https://doi.org/10.3390/su12219031>
- Post, S., Kleinen-von Königslöw, K., & Schäfer, M. S. (2019). Between Guilt and Obligation: Debating the Responsibility for Climate Change and Climate Politics in the Media. *Environmental Communication*, 13(6), 723–739. <https://doi.org/10.1080/17524032.2018.1446037>
- Shropshire, A. (2019). *HDI vs Ecology Plotly*. Github. https://github.com/as6140/hdi_v_ecology_plotly_blog
- Tiezzi, E., Pulselli, F., & Cristina, C. (2012). Biocapacity vs Ecological Footprint of world regions: A geopolitical interpretation. *Ecological Indicators*, 16, 23–30. <https://doi.org/10.1016/j.ecolind.2011.09.002>
- Trainer, F. E. (1990). Environmental significance of development theory. *Ecological Economics*, 2(4), 277–286. [https://doi.org/https://doi.org/10.1016/0921-8009\(90\)90015-M](https://doi.org/https://doi.org/10.1016/0921-8009(90)90015-M)
- Wackernagel, M., Hanscom, L., & Lin, D. (2017). Making the sustainable development goals consistent with sustainability. *Frontiers in Energy Research*, 5(JUL). <https://doi.org/10.3389/fenrg.2017.00018>