

Forest Fire

Gaizkia Adeline Atmaka ¹, Luthfi Izza Pratama ², Moethia Shakira ³, Sidi Janardhana Gatra Parahita ⁴

¹ 2501972493, *Computer Science and Mathematics, Faculty of Computer Science, Bina Nusantara University.*

² 2502034912, *Computer Science and Mathematics, Faculty of Computer Science, Bina Nusantara University.*

³ 2501987091, *Computer Science and Mathematics, Faculty of Computer Science, Bina Nusantara University.*

⁴ 2501995175, *Computer Science and Mathematics, Faculty of Computer Science, Bina Nusantara University.*

Abstract

Hutan adalah ekosistem yang kompleks di mana pohon merupakan bentuk kehidupan yang dominan. Keberadaan hutan sangat penting bagi keseimbangan ekologi, namun kebakaran hutan dapat menyebabkan kerusakan besar. Kebakaran hutan didefinisikan sebagai pembakaran

tanaman yang tidak terkendali dan tidak ditentukan dalam pengaturan alami seperti hutan, padang rumput, semak belukar atau tundra. Sistem Fire Weather Index (FWI) digunakan untuk memprediksi potensi kebakaran hutan, membantu dalam mengambil tindakan pencegahan, dan pemadaman kebakaran hutan. Tindakan preventif seperti pengelolaan lahan yang baik dan pemantauan cuaca diperlukan untuk mengurangi risiko kebakaran hutan.

Keywords: *wildfire*, Fire Weather Index (FWI), Fine Fuel Moisture Code (FFMC), Duff Moisture Code (DMC), Drought Code (DC), Initial Spread Index (ISI).

1. Overview

Hutan adalah sistem ekologi yang kompleks di mana pohon merupakan bentuk kehidupan yang dominan. Hutan adalah ekosistem alam yang paling efisien, dengan tingkat fotosintesis yang tinggi yang memengaruhi sistem tanaman dan hewan dalam serangkaian hubungan organik yang kompleks. Namun, kebakaran hutan (wildfire) dapat menyebabkan kerusakan besar pada ekosistem ini. Kebakaran hutan didefinisikan sebagai pembakaran tanaman yang tidak terkendali dan tidak ditentukan dalam pengaturan alami seperti hutan, padang rumput, semak belukar atau tundra, yang mengkonsumsi bahan bakar alami dan menyebar berdasarkan pada kondisi lingkungan.

Untuk memprediksi potensi kebakaran hutan, sistem Fire Weather Index (FWI) digunakan. Dataset yang digunakan dalam analisis ini mencantumkan empat indeks dari sistem tersebut, yaitu Fine Fuel Moisture Code (FFMC), Duff Moisture Code (DMC), Drought Code (DC), dan Initial Spread Index (ISI). Indeks-indeks ini didapatkan melalui pengukuran harian kondisi cuaca pada lokasi dan waktu tertentu. Sistem FWI ini dapat digunakan untuk memprediksi potensi kebakaran hutan dan membantu dalam mengambil tindakan pencegahan dan pemadaman kebakaran hutan.

Selain itu, kebakaran hutan dapat dipicu oleh tindakan manusia, seperti pembukaan lahan, kekeringan ekstrim, atau

dalam kasus yang jarang terjadi oleh petir. Oleh karena itu, penting untuk mengambil tindakan preventif seperti pengelolaan lahan

2. Background

Forest atau dalam bahasa Indonesia adalah hutan, merupakan sistem ekologi yang kompleks di mana pohon adalah bentuk kehidupan yang dominan dan ekosistem alam yang paling efisien, dengan tingkat fotosintesis yang tinggi yang memengaruhi sistem tanaman dan hewan dalam serangkaian hubungan organik yang kompleks. Hutan dapat berkembang dalam berbagai kondisi, dan jenis kehidupan tanah, tumbuhan, dan hewan berbeda sesuai dengan pengaruh lingkungan yang ekstrim.

Fire atau api adalah oksidasi cepat dari suatu bahan bakar dalam proses pembakaran kimia, yang melepaskan panas, cahaya, dan berbagai produk reaksi. Pada titik tertentu dalam reaksi pembakaran, yang disebut titik penyalaan, api dihasilkan. Nyala api adalah bagian api yang terlihat. Nyala api terutama terdiri dari karbondioksida, uap air, oksigen dan nitrogen. Api biasa digunakan oleh manusia dalam pertanian untuk membuka lahan, untuk memasak, menghasilkan panas dan cahaya, peleburan, penguapan, pembakaran

yang baik dan pemantauan cuaca untuk mengurangi risiko kebakaran hutan.

sampah, kremasi, dan sebagai senjata atau cara penghancuran.

Forest Fire, *wildfire*, atau biasa disebut kebakaran hutan dapat digambarkan sebagai setiap pembakaran atau pembakaran tanaman yang tidak terkendali dan tidak ditentukan dalam pengaturan alami seperti hutan, padang rumput, semak belukar atau tundra, yang mengkonsumsi bahan bakar alami dan menyebar berdasarkan pada kondisi lingkungan (misalnya, angin, topografi). Kebakaran hutan yang besar dapat menyebar dengan cepat melalui cabang pohon dengan ketinggian lebih dari 1,8 meter (6 kaki) mengakibatkan ledakan. Kebakaran hutan dapat dipicu oleh tindakan manusia, seperti pembukaan lahan, kekeringan ekstrim, atau dalam kasus yang jarang terjadi oleh petir (IRDR).

Fire Weather Index (FWI) adalah sebuah sistem indeks yang digunakan untuk memprediksi potensi kebakaran hutan. Nilai-nilai indeks tersebut didapatkan melalui pengukuran harian kondisi cuaca pada lokasi dan waktu tertentu. Dataset yang

digunakan dalam analisis ini mencantumkan empat indeks dari sistem tersebut, yaitu Fine Fuel Moisture Code (FFMC), Duff Moisture Code (DMC), Drought Code (DC), dan Initial Spread Index (ISI).

- Fine Fuel Moisture Code: mewakili kelembaban bahan bakar di bawah kanopi hutan dan di permukaan paling atas tanah.
- Duff Moisture Code: mewakili kelembaban materi organik yang diuraikan di bawah lapisan teratas.
- Drought Code: mewakili kekeringan tanah sampai lapisan bawah.
- Initial Spread Index: mewakili potensi penyebaran kebakaran dengan menggabungkan kelembaban bahan bakar di permukaan tanah dan kecepatan angin.

Rstudio merupakan cara mudah yang dilakukan untuk berinteraksi dengan

pemrograman R, hal ini dianggap sebagai Integrated Development Environment (IDE) yang menyediakan solusi untuk semua komputasi statistik dan grafis. RStudio adalah versi R yang dilengkapi dengan pengaturan *multi-pane window setup* yang menyediakan akses ke semua hal utama pada single screen (seperti source, console, environment & history, files, photos, graphs, etc).

UCI Machine Learning Repository berfungsi sebagai repository dataset untuk digunakan dalam aplikasi machine learning dan bidang-bidang terkait. UCI Machine Learning menyatukan keterampilan komputasi dan statistik serta pembelajaran mesin untuk pemecahan masalah berbasis data. Area yang berkembang pesat ini mencakup pembelajaran mendalam, analisis data berskala besar dan memiliki aplikasi dalam *e-commerce*, pencarian informasi, pemodelan bahasa alami, keuangan, bioinformatika, dan area terkait dalam kecerdasan buatan.

3. Objectives

- Mendefinisikan *forest fire* dan *fire weather index* (FWI) serta menjelaskan kaitannya.
- Mengidentifikasi kondisi lingkungan yang mempengaruhi potensi kebakaran hutan.
- Menjelaskan cara kerja sistem FWI dalam memprediksi potensi kebakaran hutan.
- Menganalisis data FWI untuk menentukan tingkat potensi kebakaran hutan pada lokasi dan waktu tertentu.
- Melihat apakah FWI dapat digunakan sebagai prediktor untuk area.

4. Methodology

Metode yang digunakan dalam penelitian ini adalah metode kuantitatif. Metode kuantitatif merupakan metode yang biasa disebut sebagai metode discovery, karena dengan menggunakan metode ini dapat ditemukan dan dikembangkan sebagai iptek baru dengan data penelitian berupa angka-angka dan analisis statistik. Data set yang dipakai oleh penulis diambil dari website UCI Machine Learning yang bernama “Forest Fires”. Berdasarkan deskripsi pada data set tersebut, diketahui bahwa lokasi hutan yang terbakar berada di wilayah timur laut negara Portugal.

Penelitian mulai dilaksanakan pada tanggal 1 Januari 2022 yang berlokasi di kediaman masing-masing peneliti. Tujuan dari penelitian ini adalah untuk memprediksi besarnya luas daerah yang terbakar. Tahapan analisis data yang dilakukan penulis adalah sebagai berikut : (1) Memfilter data yang tidak diperlukan, (2) Membuat pemodelan, (3) Melakukan uji spesifikasi parameter, (4) Melakukan uji kebaikan model, (5) Menentukan hipotesis untuk setiap parameter, dan (6) Menentukan kesimpulan.

5. Result and Discussion

Data yang dipakai untuk *article* ini adalah data yang didapatkan dari *website* UCI Machine Learning. Data yang dipakai yaitu “*forest fires*”. Tahapan analisis data yang dilakukan penulis adalah sebagai berikut : (1) Memfilter data yang tidak diperlukan, (2) Membuat pemodelan, (3) Melakukan uji spesifikasi parameter, (4) Melakukan uji kebaikan model, (5) Menentukan hipotesis untuk setiap parameter, (6) Menentukan kesimpulan.

Setelah itu membuat *dataframe* baru dimana seluruh *observation* *area* = 0 dibuang, dan *variables* yang disimpan hanya *variables* FWI (Fire Weather Index) dan *variables* *area*. Mengapa *variables* FWI tidak dibuang? Karena FWI memang sudah memperhitungkan kondisi dari cuaca berdasarkan data tersebut. Lalu mengapa koordinat dibuang datanya? Koordinat dibuang karena menurut penulis koordinat tidak terlalu penting untuk dijadikan analisis. Sehingga penulis hanya memakai *variables* FWI untuk memodelkan luas hutan yang terbakar.

Setelah membuang data yang tidak penulis perlukan untuk analisis, penulis membuat histogram dari *variables* *area* untuk melihat penyebaran datanya [fig 1].

Untuk *variables* data seperti ini penulis berekspektasi bahwa penyebarannya akan mendekati distribusi normal, tetapi setelah melihat dari histogram yang telah penulis buat bahwa penyebaran dari *variables* *area* tidak mendekati distribusi normal. Berdasarkan hasil dari histogram yang telah dibuat bahwa data dari *variables* *area* lebih cenderung condong ke nilai - nilai yang lebih kecil. Karena hasil histogram bahwa nilainya condong lebih kecil, sehingga data ini tidak bisa memakai “Linear Regression”.

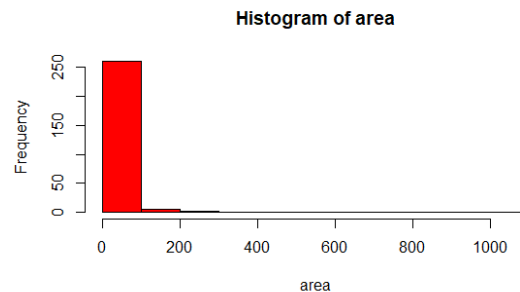


fig 1

Karena tidak bisa menggunakan linear regression akhirnya penulis mengikuti sumber data dari *website* UCI Machine Learning dan menemukan bahwa penulis dapat menggunakan “Log Transform” sebagai pengganti dari linear regression. Sehingga *variables* *area*, penulis ubah menjadi $\ln(\text{area})$. Setelah penulis coba membuat histogram dengan menggunakan

$\ln(\text{area})$ hasil dari histogram tersebut yaitu penyebaran variables area menjadi mendekati distribusi normal.

Lalu penulis membuat Q-Q Plot untuk area data [fig 2]. Fungsi dari grafik qq norm yaitu untuk membuat konfirmasi lebih kuat bahwa data yang telah di transform telah mengikuti distribusi normal. Setelah itu, penulis membuat data frame baru dengan menambahkan $\ln(\text{area})$ dan membuang variables area yang biasa atau tidak diperlukan [fig 3.a - 3.b].

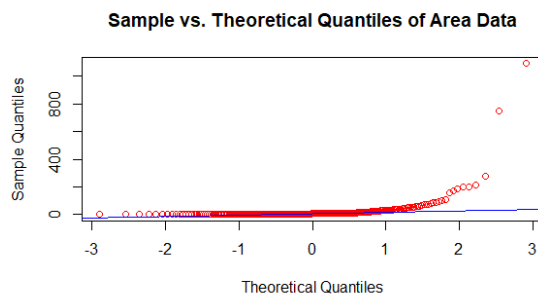


fig 2

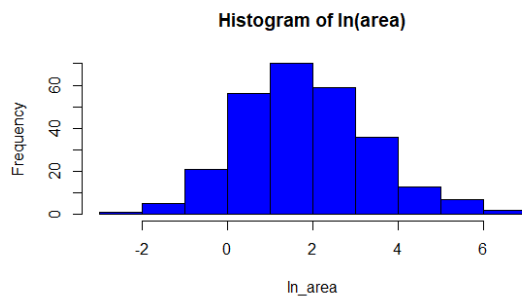


fig 3.a

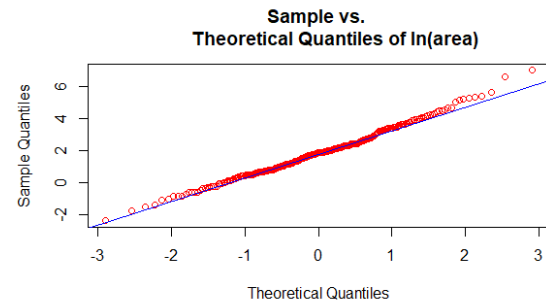


fig 3.b

Setelah membuat data frame baru penulis membuat pemodelan memakai linear regression. Berdasarkan hasil dari pemodelan tersebut bahwa data banyak yang berubah $\ln(\text{area})$ berubah menjadi variables terikat dan sisa dari variables datanya itu sebagai variables bebas. Sebagai pembuktiannya penulis membuat grafik untuk melihat hasilnya menggunakan Added-Variable Plots (avplots) [fig 4].

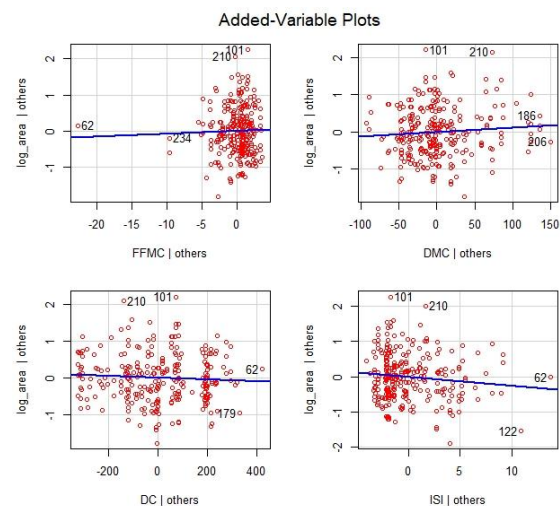


fig 4

Setelah membuat grafik menggunakan av plots, langkah selanjutnya yaitu men-generate variables yang isinya berupa *variables* residuals. Apa itu residuals? Residuals merupakan jarak masing - masing data dari model regresi. Setelah generate *variables* penulis membuat plot residuals vs nilai - nilai dari model linearnya [fig 5]. Jika pada plot tersebut terdapat penyebaran di sekitar $y = 0$, maka dapat dikonfirmasi bahwa memang ada hubungan antara variabel bebas dan variabel terikat.

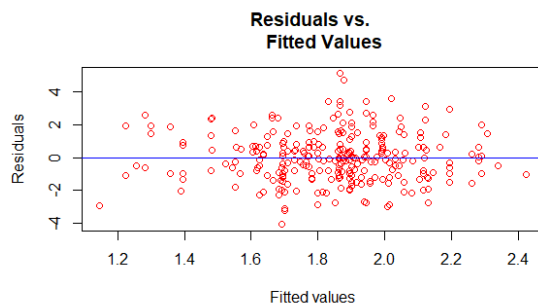


fig 5

Setelah itu membuat Q-Q Plot untuk membuat residual menggunakan qqnorm [fig 6]. Jika pada pemodelan menggunakan qqnorm titik-titiknya mengumpul disekitar garis, maka dapat dikatakan asumsi linearitasnya valid. Setelah itu membuat density plot menggunakan residuals [fig 7]. Density plot ini digunakan untuk melihat

apakah penyebarannya kurang lebih mengikuti distribusi normal.

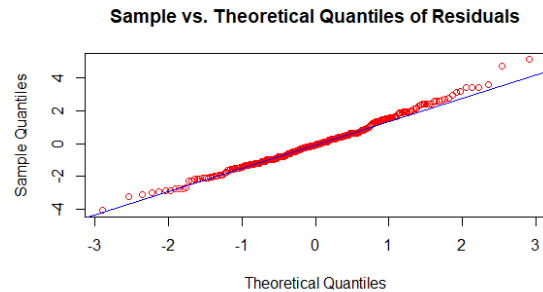


fig 6

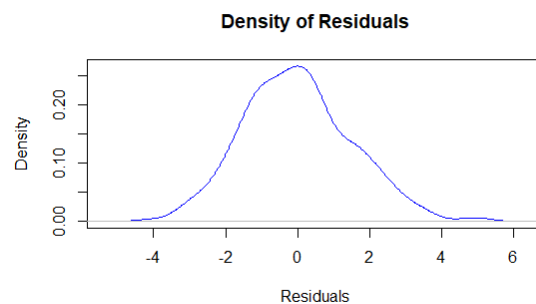


fig 7

Langkah terakhir yaitu menghitung wald test untuk membuat modelnya. Apa itu wald test? Wald test merupakan tes yang dipakai untuk melihat variables bebas yang dipakai apakah signifikan atau tidak, wald test digunakan untuk memperhitungkan keberadaan dari variable - variabel bebas yang lain sehingga jika ada variables yang tidak terlalu penting dapat dibuang. Untuk itu dapat dilihat pada p-value. Penulis menggunakan taraf signifikansi $\alpha = 0,05$ sehingga jika $p\text{-value} < 0.05$ maka dapat

dikatakan bahwa variabel tersebut signifikan. Jika $p\text{-value} > 0.05$ maka dapat kita anggap bahwa variabel tersebut aman untuk dibuang.

Setelah di uji variables nya satu - satu hasil dari pengujian tersebut menunjukkan bahwa $p\text{-value} > 0.05$ [fig 8.a - 8.d], dapat disimpulkan bahwa variables dari data ini merupakan variables tidak signifikan. Setelah menguji variables satu - satu, penulis juga menguji variables nya secara bersamaan [fig 9]. Jika variabelnya diuji secara bersamaan maka hasil dari $p\text{-value}$ menjadi sangat kecil. Sehingga dapat diambil berdasarkan hasil pengujian menggunakan wald test secara bersamaan yang berarti variabelnya signifikan jika dan hanya jika semuanya dipakai secara bersamaan.

Saat penulis membuat pemodelan dengan regresi linear didapatkan hasil yang tidak signifikan, yaitu $p\text{-value} > 0.05$ dapat dilihat pada bagian `summary(model)`. Sehingga ada asumsi linearitas yang tidak terpenuhi. Saat dibuat plotnya terlihat seperti terpenuhi walau tidak sempurna. Tetapi saat dibuat regresi linear antara variabel bebasnya satu-satu didapatkan hasil yang signifikan, yaitu $p\text{-value} < 0.05$. Hal ini menunjukkan bahwa variabel bebasnya tidak independen satu sama lain, ada korelasi di

antara mereka. Sehingga kurang tepat jika dibuat model menggunakan regresi linear.

Penulis memakai uji collinearity antar variabel bebasnya karena jika hanya menggunakan variabel FWI saja, variabel tersebut bukan prediktor yang cukup baik untuk memprediksi area hutan yang terbakar.

```
> wald.test(Sigma = vcov(model), b = coef(model), Terms = 1)
wald test:
-----
Chi-squared test:
X2 = 0.048, df = 1, P(> X2) = 0.83
```

fig 8.a

```
> wald.test(Sigma = vcov(model), b = coef(model), Terms = 2)
wald test:
-----
Chi-squared test:
X2 = 0.23, df = 1, P(> X2) = 0.63
```

fig 8.b

```
> wald.test(Sigma = vcov(model), b = coef(model), Terms = 3)
wald test:
-----
Chi-squared test:
X2 = 1.9, df = 1, P(> X2) = 0.17
```

fig 8.c

```
> wald.test(Sigma = vcov(model), b = coef(model), Terms = 4)
wald test:
-----
Chi-squared test:
X2 = 1.1, df = 1, P(> X2) = 0.3
```

fig 8.d

```
> wald.test(Sigma = vcov(model), b = coef(model), Terms = 1:4)
wald test:
-----
Chi-squared test:
X2 = 99.4, df = 4, P(> X2) = 0.0
```

fig 9

6. Conclusion

Dapat disimpulkan berdasarkan hasil dari analisis data menggunakan data “forest fires” bahwa data yang dipakai tidak bisa menggunakan “Linear Regression” karena penyebaran dari variables area tidak mendekati normal. Sehingga, penulis menggunakan “Log Transform” dengan menambahkan $\ln(\text{area})$ dan hasil histogram menjadi mendekati distribusi normal, sedangkan pemodelan Q-Q Plot area data untuk mengkonfirmasi data di transform mengikuti distribusi normal. Hasil dari pemodelan variable baru menunjukkan $\ln(\text{area})$ berubah menjadi variables terikat dan sisa dari variables datanya itu sebagai variables bebas. Sedangkan hasil dari avplots yaitu terjadi penyebaran di sekitar $y = 0$, sehingga terdapat hubungan antara

variabel bebas dan variabel terikat. Untuk pemodelan Q-Q Plot menggunakan qqnorm titik mengumpul disekitar garis, sehingga diasumsikan linearitasnya valid. Setelah itu density plot menggunakan residuals. Langkah terakhir menguji data menggunakan wald test, hasil wald test dari pengujian satu-satu yaitu $p\text{-value} > 0.05$ sehingga variables dari data ini merupakan variables tidak signifikan. Sedangkan jika data diuji secara bersamaan menghasilkan $p\text{-value}$ sangat kecil, sehingga diambil berdasarkan hasil pengujian menggunakan wald test secara bersamaan yang berarti variabelnya signifikan jika dan hanya jika semuanya dipakai secara bersamaan.

7. Reference

- *Forest Fires Data Set*. (2008, February 29). UCI Machine Learning Repository. Retrieved January 19, 2023, from <https://archive.ics.uci.edu/ml/datasets/Forest+Fires>
- *Forest | Definition, Ecology, Types, Trees, Examples, & Facts*. (2023, January 6). Encyclopedia Britannica. Retrieved January 19, 2023, from <https://www.britannica.com/science/forest>

- *Fire*. (n.d.). Wikipedia. Retrieved January 19, 2023, from <https://en.wikipedia.org/wiki/Fire>
- *What is RStudio? Complete Guide to What is RStudio*. (n.d.). eduCBA. Retrieved January 19, 2023, from <https://www.educba.com/what-is-rstudio>
- Kane, J. (n.d.). *Forest fire | Definition, Description, Ecology, & Facts*. Encyclopedia Britannica. Retrieved January 19, 2023, from <https://www.britannica.com/science/forest-fire>
- *Forest Fire | UN-SPIDER Knowledge Portal*. (n.d.). UN-Spider. Retrieved January 19, 2023, from <https://www.un-spider.org/category/disaster-type/forest-fire>
- *Machine Learning MSc | Prospective Students Graduate*. (n.d.). University College London. Retrieved January 19, 2023, from <https://www.ucl.ac.uk/prospective-students/graduate/taught-degrees/machine-learning-msc>
- *Data Science and Machine Learning MSc | Prospective Students Graduate*. (n.d.). University College London. Retrieved January 19, 2023, from <https://www.ucl.ac.uk/prospective-students/graduate/taught-degrees/data-science-and-machine-learning-msc>
- *Penyebab Terjadinya Kebakaran Hutan – TBNews Polda Kepri*. (2021, September 15). TRIBATANews POLDA KEPRI. Retrieved January 19, 2023, from <https://tribatanews.kepri.polri.go.id/2021/09/15/penyebab-terjadinya-kebakaran-hutan/>
- *Home*. (n.d.). YouTube. Retrieved January 19, 2023, from <https://www.nwcg.gov/publications/pms437/cffdrs/fire-weather-index-syst>