**LAPORAN EKSPERIMEN**

Pengantar Data Science

**Explorasi Dataset Forest Fires**

forestfire.csv

**Disusun Oleh**

William Surya Wijaya – 6182001004

Kae Zarren Luvoma H – 6182001004

Ester Oka Nugraheni – 6182001042

**Tanggal**

Sabtu, 26 November 2022

**Jurusan Informatika**

UNIVERSITAS KATOLIK PARAHYANGAN

1. Tujuan Eksperimen

Tujuan dari eksperimen dari dataset forestfirest.csv adalah untuk mengetahui dan mengenali data berserta hubungan antar antribut untuk mengambil kesimpulan mengenai faktor faktor yang berhubungan dengan luas kebakaran hutan dan intensitas kebakaran hutan.

1. Deskripsi Dataset, Penyiapan Data, dan Hipotesis

Dataset forestfires.csv memiliki dimensi (517, 13)

Berikut adalah deskripsi dataset berupa nama atribut, keterangan berserta nilai minimal dan maksimal dari masing masing atribut.

|  |  |  |  |  |
| --- | --- | --- | --- | --- |
| No | Atribut | Keterangan | Min | Max |
| 1. | X | Koordinat x dari lokasi kebakaran hutan di dalam lingkup area Montesinho Park. | 1 | 9 |
| 2. | Y | Koordinat y dari lokasi kebakaran hutan di dalam lingkup area Montesinho Park. | 2 | 9 |
| 3. | Month | Bulan terjadinya kebakaran hutan (January – December) dengan format 3 first letters. | Jan | Dec |
| 4. | Day | Hari terjadinya kebarakan hutan (Monday – Sunday) dengan format 3 first letters. | Mon | Sun |
| 5. | FFMC | Fine Fuel Moisture Code - Index tingkat potensi terjadinya kebakaran yang diukur dari kelembaban serasah dan bahan halus mudah terbakar di sekitar. | 18.7 | 96.2 |
| 6. | DMC | Duff Moisture Code - Index tingkat potensi terjadinya kebakaran yang diukur dari kelembaban rata rata lapisan tanah organik dan bahan berukuran sedang yang mudah terbakar di sekitar. | 1.1 | 291.3 |
| 7. | DC | Drought Code – Index tingkat potensi terjadinya kebakaran yang diukur dari kelembaban lapisan tanah organik padat dan bahan berukuran besar (bongkahan) yang mudah terbakar di sekitar dan potensi terjadinya kabut asap. | 7.9 | 860.6 |
| 8. | ISI | Initial Spread Index - Index perkiraan penyebaran api yang diukur dari FFMC dan pergerakan angin. | 0.0 | 56.1 |
| 9. | Temp | Temperatur sekitar dalam satuan derajat celcius. | 2.2 | 33.3 |
| 10. | RH | Kelembaban sekitar dalam satuan %. | 15 | 100 |
| 11. | Wind | Kecepatan angin dalam satuan km/h. | 0.4 | 9.4 |
| 12. | Rain | Kepadatan hujan dengan satuan mm/m2. | 0.0 | 64 |
| 13. | Area | Luas area kebakaran hutan dengan satuan ha. | 0.0 | 1090.84 |

Dataset forestfires.csv tidak memiliki nilai NaN.

Hipotesis dan Hipotesa

1. Apa faktor yang berpengaruh secara signifikan terhadap atribut terget (area) ?
2. Pada bulan mana sajakah yang memiliki tingkat intensitas kebakaran hutan tinggi ?
3. Pada bulan mana sajakah yang memiliki luas area kebakaran hutan tinggi ?
4. Pada bulan mana sajakah yang memiliki tingkat temperatur tinggi ?
5. Penelitian

Mengetahui atribut atribut yang memiliki pengaruh signifikan baik hubungan positif dan hubungan negatif terhadap atribut target (area) dengan menggunakan metode Heat-Map.

A picture containing treemap chart

Description automatically generated

Visualisasi hubungan atribut atribut prediktor dengan atribut target tidak menunjukan hubungan korelasi yang kuat. Berikut adalah beberapa atribut dengan nilai korelasi tertinggi.