**I. PENDAHULUAN**

Salah satu masalah lingkungan yang utama adalah terjadinya kebakaran hutan (juga disebut kebakaran hutan liar), yang mempengaruhi pelestarian hutan, menimbulkan kerusakan ekonomi dan ekologi, serta menyebabkan penderitaan manusia.

 Laporan ini bertujuan untuk mengembangkan model klasifikasi yang dapat memprediksi kemungkinan terjadinya kebakaran hutan berdasarkan parameter cuaca seperti suhu, kelembaban, kecepatan angin, dan curah hujan.

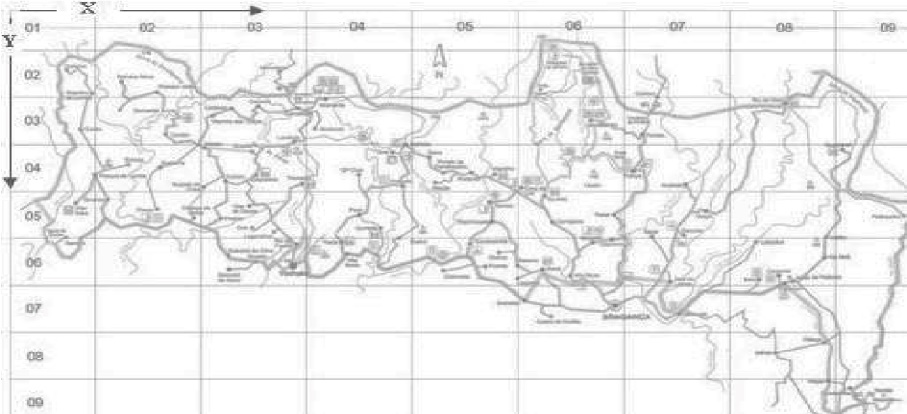
**II. DESKRIPSI DATASET**

Dataset yang digunakan berasal dari UCI Machine Learning Repository Sumber dataset (<https://archive.ics.uci.edu/dataset/162/forest+fires>) .

1. **Data Kebakaran Hutan**

Data peeniltian ini diambil dari data kebakaran hutan taman alam Monteshino dari wilayah timur laut di portugal.

Terdapat 13 variabel pada sumber dataset. Diantaranya nilai koordinat spasial pada sumbu X (1-9) dan nilai pada sumbu Y (2-9) yang tertera pada peta taman Monteshino **Gambar 1**.



**Gambar 1.** PetaTaman NasionalMonteshino

Terdapat juga variabel bulan (month), hari (day), Fine Fuel Moisture Code (FFMC), Duff Moisture Code (DMC), Drought Code (DC), Initial Spread Index (ISI), data ini merupakan Forest Fire Weather Index (FWI). FWI adalah sistem Kanada untuk menilai bahaya kebakaran dan mencakup enam komponen tersebut. temp (suhu dalam °C), RH (kelembapan Relatif (%)), *wind* (kecepatan angin (km/jam)), *rain* (curah hujan (mm/m2)) dan area (ditransformasi dengan variabel biner fire (1 jika area > 0 , 0 jika area <= 0).

jumlah data yang diambil sebanyak 517 pada periode Januari tahun 2000 sampai dengan Desember 2003 diwilayah timur portugal.

**2. Preprosesing Data**

Pada tahap awal preprosesing data yaitu dengan mengidentifikasi nilai yang kosong atau null. Terdapat 12 nilai yang hilang pada kolom rain, dan diisi dengan nilai median dan 0.0 mm.

Perubahan variable area menjadi fire agar memudahkan dalam menentukan variabel target (Y), dengan kriteria Fire = 0 jika area = 0 artinya tidak ada kebakaran, dan fire = 1 jika area > 0 artiya terjadi kebakaran.

Variable lainnya yaitu temp(suhu), RH (Relative Humidity/kelembapan), Wind (kecepatan angin), Rain (curah hujan).

**III. METODE PENELITIAN**

Metode yang digunakan yaitu model K-Nearest Neighbors (KNN). parameter nilai k (n neighbors) yang digunakan yaitu 5. Pembagian dataset menjadi data latih 362 sampel (70%) dan data uji 155 sampel (30%). Kemudian dilakukan normalisasi data dengan menggunakan metode StandarScaler untuk menyamakan sekala fitur.

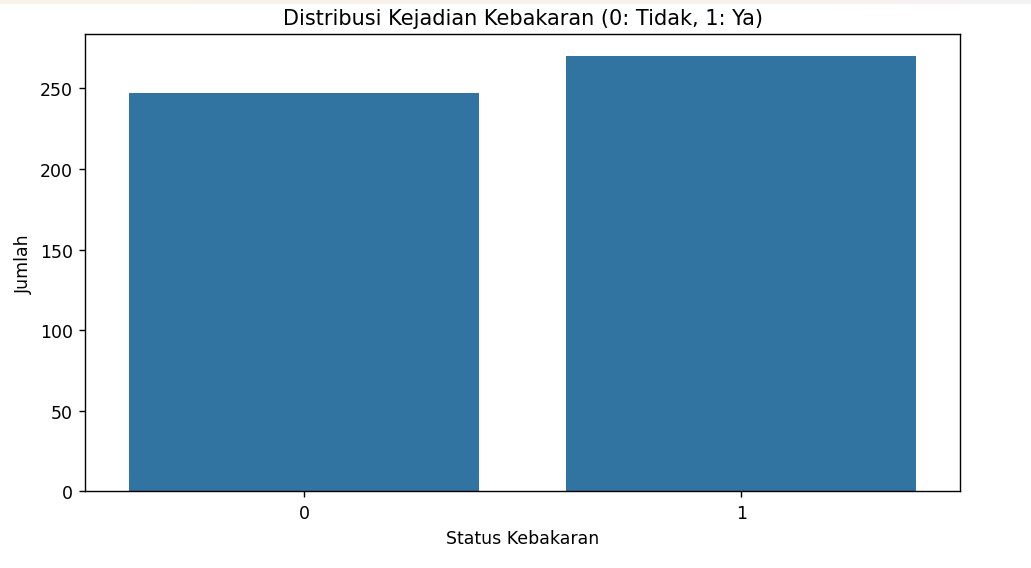
Sofware pengolahan data yang digunakan menggunakan bahasa pemograman python dengan tools yang digunakan adalah visual studio code.

**IV. HASIL PENELITIAN DAN PEMBAHASAN**

* 1. **Korelasi antar variabel dengan kebakaran**

**4.1.1 Distribusi Area Target**

Hasil pengolahan data untuk distribusi kejadian kebakaran dapat diketahui melalui grafik dibawah ini.



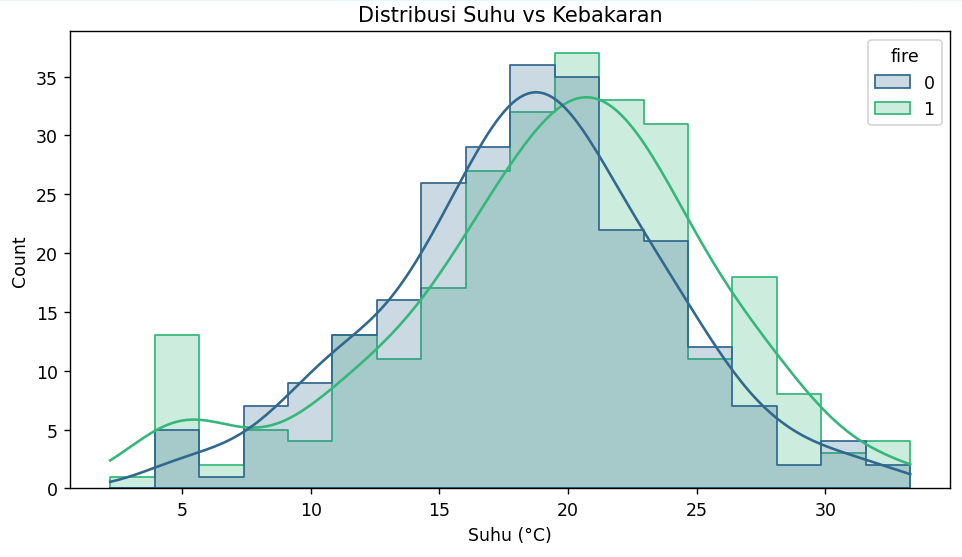
**Gambar 2**. Distribusi Kejadian kebakaran

Distribusi kelas pada grafik diaatas diketahui area tidak mengalami kebakaran (area = 0) sebanyak 242 data atau 43.5%, dan area yang mengalami kebakaran sebanyak 292 (56.5%) data mengalami kebakaran (area > 0).

Pola kebakaran lebih sering terjadi pada suhu > 15°C dengan suhu puncak di 20-25°C. Intesitas kelembaban dominan ketika kebakaran pada nilai < 60%, kecepatan angin pada kisaran > 4 km/jam meningkatkan resiko kebakaran. Korelasi suhu dengan kebakaran berkorelasi positif (0.38) dan (-0.27).

**4.1.2 Distribusi Suhu (°C) terhadap kejadian kebakaran (fire),**

Berikut adalah hasil dari pengolahan data menjadi grafik histogram suhu dengan kebakaran pada taman nasional monteshino.



Gambar 3. Distribusi suhu dengan kebakaran

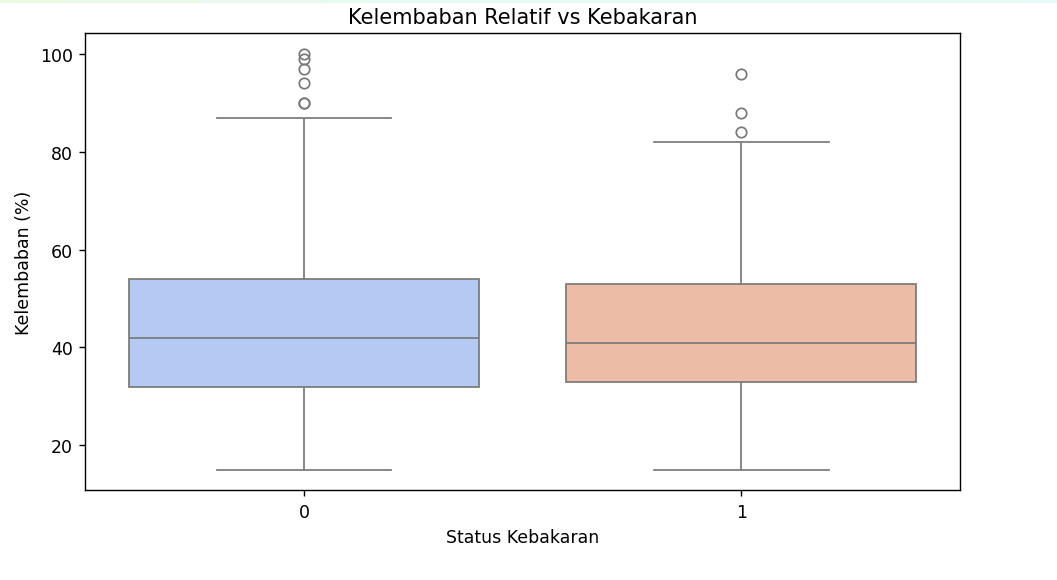
Dari grafik gambar 3 ada dua histogram warna biru mewakili tidak terjadi kebakaran (fire = 0) dan histogram warna hijau mewakili terjadinya kebakara (fire =1).

Distribusi pada saat tidak terjadi kebakaran (fire=0) menyerupai distribusi normal dengan puncak sekitar 19-20 **°**C, disuhu <10**°**C dan >28**°**C jumlah kasus tidak terjadinya kebakaran menurun artinya saat tidak terjadi kebakaran suhu umunya berkisar di antara 15**°**C hingga 24**°**C.

Distribusi saat terjadi kebakaran (fire=1) menyerupai kurva normal, dengan puncuk kejadian 20-23**°**C, disuhu rendah <10 **°**C sangat sedikit terjadi kebakaran, namun pada suhu tinggi 24**°**C masih tinggi. Artinya menunjukan korelasi positif antara suhu tinggi dengan kebakaran.

**4.1.3 Distribusi variabel kelembaban relatif dengan kebakaran**

Berikut hasil pengolahan data variabel kelembaban dengan kebakaran jika dalam grafik boxplot untuk terjadinya pola kebakaran dengan variabel kelemababan

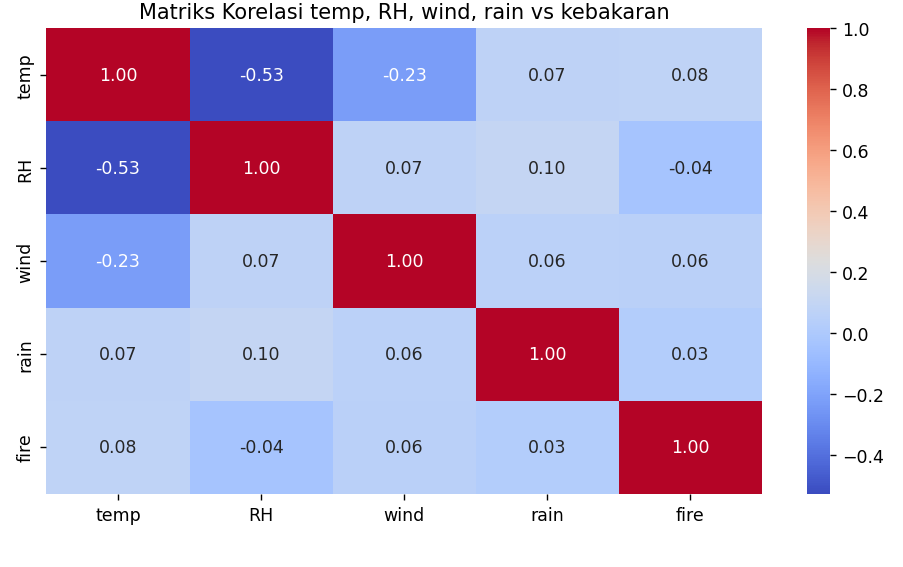


**Gambar 4**. Kelembaban Relatif dengan kebakaran

Dari gambar grafik boxplot nilai tengah box pada tidak terjadinya kebakaran (fire=0) berada dikisaran kelembaban 42%. Dan kasus tidak terjadinya kebakaran berada pada kisaran kelembaban 30-50%.

Kasus terjadinya kebakaran berada nilai tengah box plot kisaran kelembaban 40%, kasus terjadinya kebakaran berada dikisaran 35-45%, terjadinya kebakaran dengan tidak berbeda jauh dengan tidak terjadinya kebakaran, jadi kelembaban 30-55% merupakan zona dimana kebakaran paling sering maupun paling tidak terjadi dilihat dari variabel kelembaban.

Untuk lebih lengkapnya hubungan variabel lainnya dengan kebakaran bisa dilihat dari gambar 5 matrik korelasi.

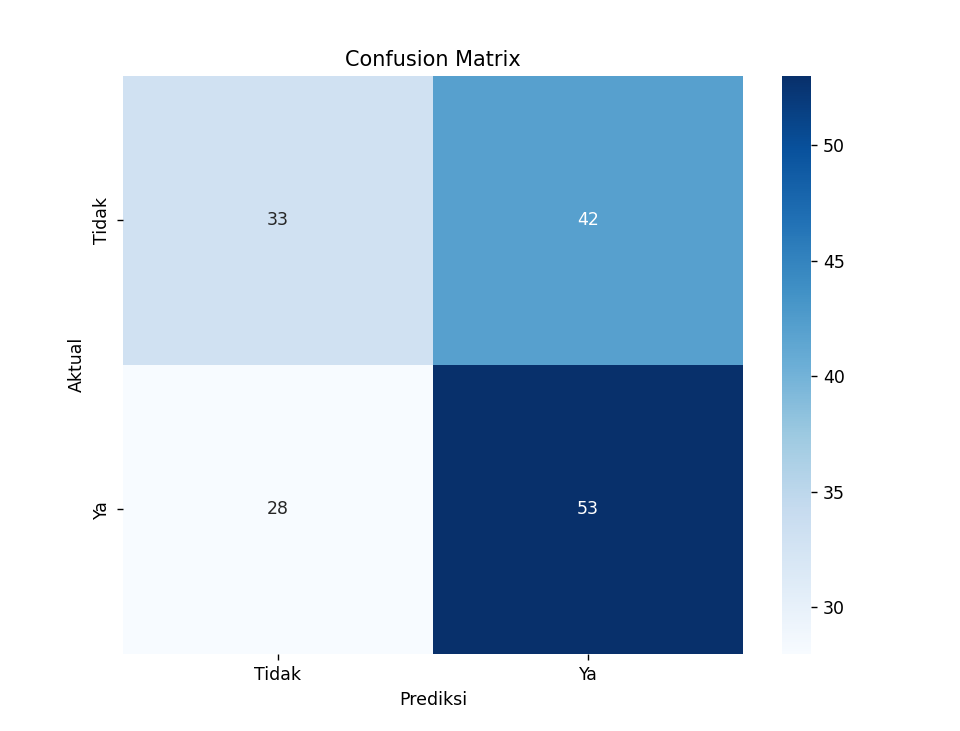


**Gambar 5.** Matrix korelasi Temp (suhu), RH (kelembaban), Rain (hujan) dengan kebakaran

Untuk korelasi temp (suhu) terhadap terjadinya kebakaran atau tidak terjadinya kebakaran sebesar 0.08 (8%), RH(kelembaban) -0.04 (4%) berkorelasi negatif artinya semakin tinggi kelembaban maka terjadinya kebakaran semakin kecil, wind 0.06 (6%), rain (curah hujan) 0.03 (3%). variable temp paling tinggi terhadap variabel lainnya, kemungkinan suhu lebih berpengaruh terhadap terjadi atau tidaknya kebakaran.

Hasil prediksi model

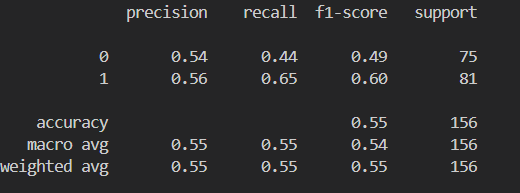
Grafik confusion matrik yang didapat dalam pemodelan menggunakan KNN seperti berikut.



Gambar 6. Confusion Matrix model

Dari confusion matrix didapatkan model memprediksi terjadinya kebakaran (True Positif) sebesar 53 dengan (false positif) sebesar 42, artinya model masih memprediksi terjadi kebakaran padahal aktualnya tidak terjadi kebakaran. Model memprediksi tidak terjadi kebakaran (false negatif )padahal masih ada kebakaran sebesar 28 dan model memprediksi tidak terjadi kebakaran (true Negatif ) sebesar 33.

Dari nilai confusion matriks diatas dengan nilai TP 53, TN 33, FP 42 dan FN 28 didapat nilai precision, recall, dan f1 score sebagai berikut



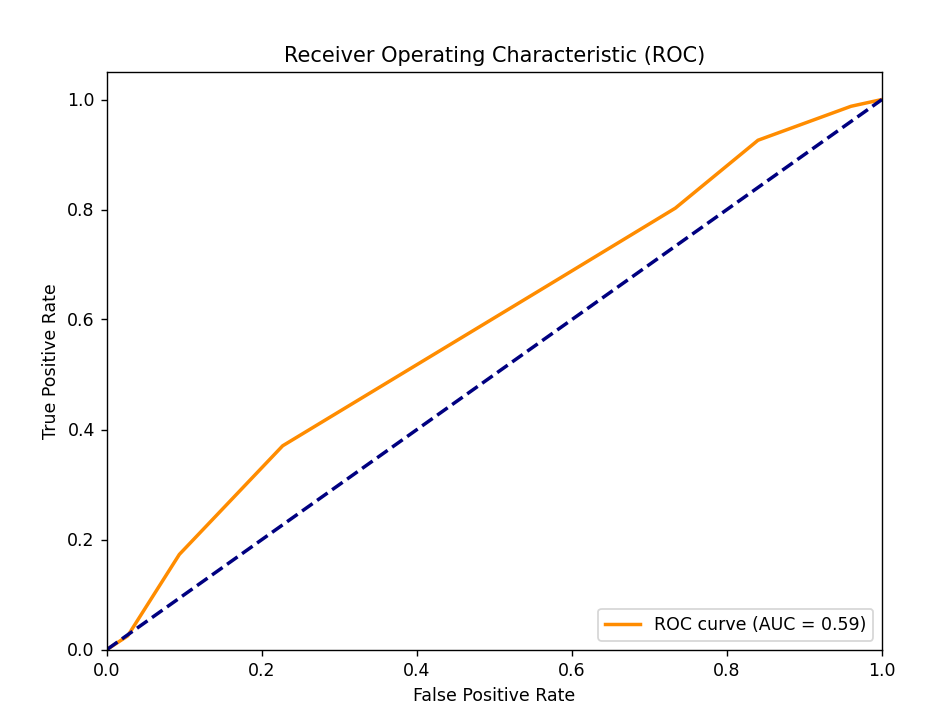
**Gambar 6.** Hasil precision, recall, f1 score

Model mampu predikis kejadian kebakaran dengan presisi 56% (precesion 1) dan tidak terjadi kebakaran dengan 54% (precision 2). Dengan akurasi model mencapai akurasi 54%.

Akurasi model yang didapat 55% artinya model tidak terlalu baik dalam memprediksi, dilihat dari segi error prediksi yang masih muncul.

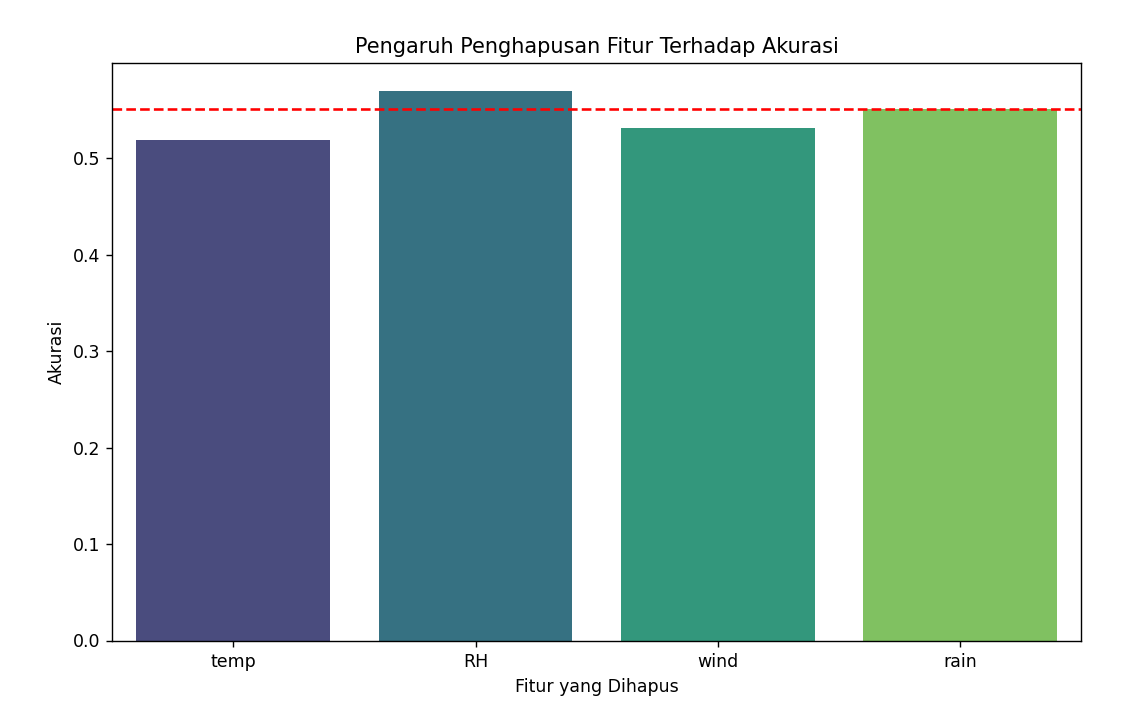
Recall kebakaran jauh lebih tinggi 65% (0.65 ) dari recall tidak terjadi kebakaran 54% (0.54) artinya model lebih sensitif terhadap kasus kebakaran.

precesion kelas 1 artinya kebakaran yang terjadi sebanyak 56% dari semua predikisi kebakaran dan sisanya 54%. Pada nilai recall kelas 1 sebesar 65%



**Gambar 7.** ROC

Nilai AUC 0.59 artinya model kurang baik dalam memprediksi kebakaran, nilai mendekati 1 maka model termasuk baik, dan mendekati 0.5 artinya model cenderung tidak baik dalam memprediksi, dan nilai yang didapat dalam pengolahan data menggunakan model KNN ini masih kurang baik.



**V. KESIMPULAN DAN SARAN**

Dari nilai

**DAFTAR PUSTAKA**

Memuat sumber-sumber yang diacu di dalam penulisan artikel, hanya sumber-sumber yang digunakan yang dimuat dalam daftar pustaka. Referensi bisa berasal dari buku, jurnal ataupun prosiding seminar serta web.

Penulisan naskah dan sitasi yang diacu dalam naskah ini disarankan format penulisan menggunakan style IEEE.

[1] Penulis 1, Penulis 2 dst. (Nama belakang, nama depan). Tahun publikasi. Judul Buku cetak miring. Edisi, Penerbit. Tempat Publikasi.

Contoh: O’Brien, J.A. dan. J.M. Marakas. 2011. *Management Information Systems*. Edisi 10. McGraw-Hill. New York-USA.

[2] Penulis 1, Penulis 2 dan seterusnya, (Nama belakang, nama depan). Tahun publikasi. Judul artikel. Nama Jurnal Cetak Miring. Vol. Nomor. Rentang Halaman.

Contoh: Sadalia, I., Syahyunan. 2016. Financial Management Behavior and Financial Distress on small medium enterprise in Indonesia., *International Journal of Applied Business and Economic Research*, Vol. 14, No. 11, (2016): 7409-7416.

[3] Penulis 1, Penulis 2 dst, (Nama belakang, nama depan disingkat). Tahun publikasi. Judul artikel. Nama Konferensi. Tanggal, Bulan dan Tahun, Kota, Negara. Halaman.

Contoh: Michael, R. 2011. Integrating innovation into enterprise architecture management. Proceeding on Tenth International Conference on Wirt-schafts Informatik. 16-18 February 2011, Zurich, Swis. Hal. 776-786.