Ujian Akhir Semester - Sistem Keamanan Cerdas

Malware Detection Using Machine Learning and Data Mining

Kelompok 1

Fathan Abdul Shodiq (1301170261)  
Fithroh Hito Naruhodo (1301174437)  
Hatami Ra’is Bukhari (1301154303)   
Shaina Maulida Safira (1301174356)

# Pendahuluan

# Studi Literatur

**Malware Analysis and Detection Using Data Mining and Machine Learning Classification**

*Paper* yang di tulis oleh Chowdhury ini mengusulkan classifier framework baru untuk mendeteksi malware dengan menggunakan gabungan dari Binary Associative Memory (BAM) dan Multilayer Perceptron (MLP) neural network. Classifier framework ini unggul dalam mendeteksi 52.185 data set yang berisi malware dan file yang terinfeksi malware dengan akurasi 98.6%, sementara dengan classifier framework lainnya seperti Naïve Bayes, J48, Random forest, dan SVM hanya mendapatkan hasil akurasi sebesar 88.5%, 91.1%, 94.8%, dan 95.7%. Hal ini membuktikan bahwa classifier network yang di ajukan penulis dapat mendeteksi malware lebih baik. [1]

**ClaMP**

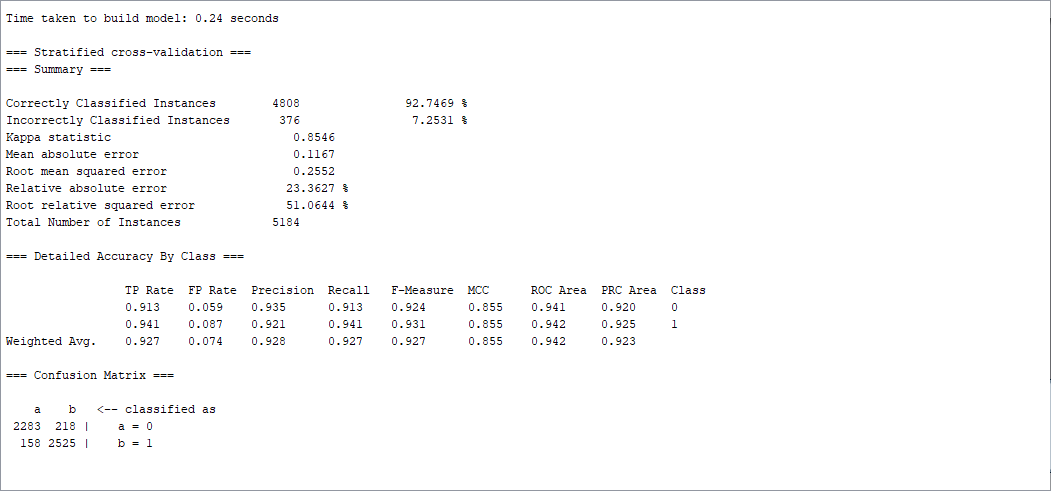
ClaMP (Classification of Malware with PE headers) merupakan github repository yang di upload oleh username urwithajit. Repository ini berisi tentang Malware classifier dataset yang di buat dengan header fields [2]. Pada tugas besar ini, kami hanya menggunakan dataset yang ada di repository dengan nama “ClaMP\_Raw-5184.arff” yang berisi 5184 sample dengan malware sebanyak 2683 dan file yang terinfeksi sebanyak 2501. Alasan kami menggunakan dataset dari repository ini adalah dataset yang ada pada paper rujukan [1], url sudah tidak bisa lagi di akses, sehigga kami harus mencari dataset baru. Hasil akan berbeda jauh dari paper rujukan [1] karena pada paper rujukan dataset ada sebanyak 52.185 sampel, sementara dataset yang kami gunakan pada tubes ini hanya 5184 sampel.

# Data Preparation

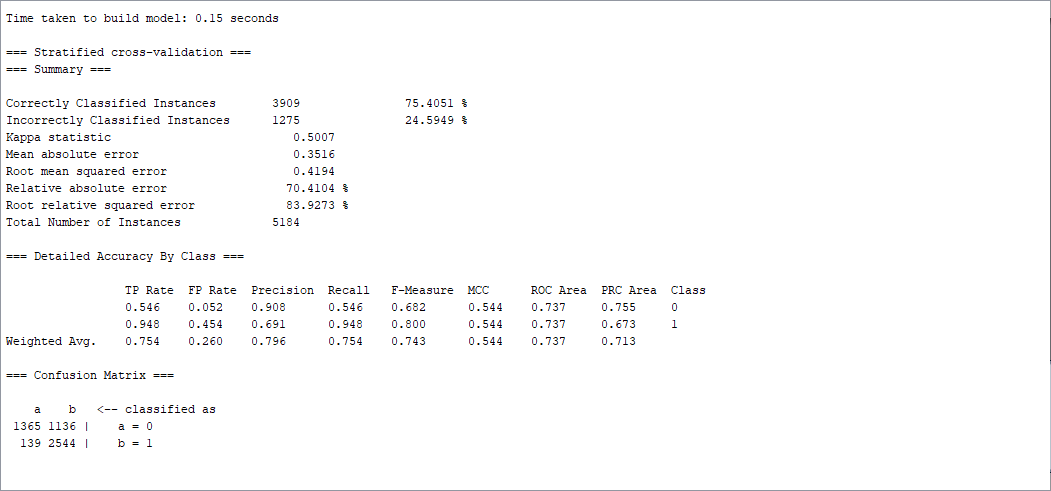
ClaMP\_Raw-5184.arff berisi sebanyak 5184 sampel (2683 malware + 2501 benign) dan memiliki 55 features (55 raw features). Pada tugas besar ini ClaMP\_Raw-5184.arff di simpan menjadi dataset-5184.arff dengan isi yang sama.

# Hasil

**Hoeffding Tree**

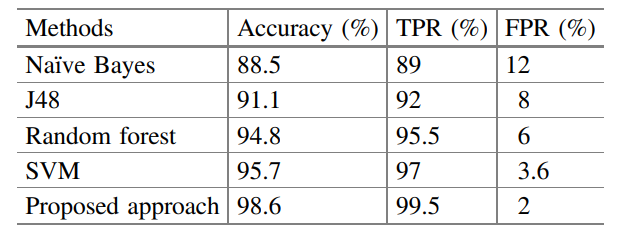


**Decision Stump**



# Analisis

Hasil dari paper rujukan



Hasil dari tugas akhir

[ table isi sama kaya di atas]

# Kesimpulan

# Referensi

|  |  |
| --- | --- |
| [1] | M. Chowdhury, A. Rahman dan R. Islam, “Malware Analysis and Detection Using Data,” dalam *International Conference on Applications and Techniques in Cyber Security and Intelligence*, 2018. |
| [2] | urwithajit9, “Github - ClaMP (Classification of Malware with PE headers),” [Online]. Available: https://github.com/urwithajit9/ClaMP. [Diakses 9 1 2021]. |

Repository

[hatamirais/tubes-sistem-keamanan-cerdas (github.com)](https://github.com/hatamirais/tubes-sistem-keamanan-cerdas)