**Cybersecurity Analysis Project**

**Overview**

The aim of the project is to analyze the cybersecurity dataset file to obtain the results and analyze- the heatmaps, attack signatures and perform classification models.

**Dataset Description**

The data set I used can be found at - <https://www.kaggle.com/datasets/teamincribo/cyber-security-attacks?select=cybersecurity_attacks.csv>

This dataset has, many columns such as -

|  |  |  |  |  |  |  |
| --- | --- | --- | --- | --- | --- | --- |
| **Timestamp** | Source IP Address | Destination IP Address | Source Port | Destination Port | Protocol | Packet Length |

**Requirements**

List of the required libraries and tools needed to run your analysis

* Python 3.x
* Pandas
* NumPy
* Seaborn
* Matplotlib
* Scikit-learn

**Installation**

pip install pandas numpy seaborn matplotlib scikit-learn

**File Structure**

The structure of my project directory

* **cybersecurity\_attacks.csv**: The dataset file.
* **skpagad\_project\_file.ipynb**: Jupyter notebook containing the analysis.

**Analysis Overview(present in the Document description in detail)**

Consists of a high-level overview of the analysis performed, including:

* Data cleaning and preprocessing steps.
* Descriptive statistics and visualizations.
* Heatmaps of attack types by hour, day of the week, and month.
* Count plots for various categorical features.
* Top locations for different types of attacks.

**Key Insights**

The Decision Tree model has the highest accuracy at 33.3%, but this is only marginally

better than the KNN and Random Forest models, which have accuracies of 32.4% and 32.8%, respectively. The closeness of these accuracy scores suggests that none of the models is significantly outperforming the others.

**Model Building(present in detail in the document)**

**Conclusion**

In conclusion, the accuracy scores suggest that further investigation and refinement of the models and data are required for better performance. Adjustments in the modeling approach, deeper data analysis, and exploring different evaluation metrics are recommended steps forwar

**Contact**

Shruti Krishna Pagadala

[SKPAGAD@BU.EDU](mailto:SKPAGAD@BU.EDU)