Humanised AI for Emotion Based Threat Prediction on the Dark Web

Ananya Sharma[1], Anushka Raj[2], Rishika Solanki[3]
[1,2,3] *Department of Computer Science and Engineering, Rajasthan College of Engineering for Women, Jaipur*

Abstract- A place on internet that is encrypted and anonymous is known as dark web. Dark web is like a untraceable hidden layer of the internet used to store and access the confidential information. Dark was originally built by U.S. Department of Defense to allow for a secure and anonymous communication among intelligence agencies and government workers. But there are number of incidents which reported the misuse of this platform for conducting the criminal and illegal activities in a hidden manner such as illegal weapons distribution, human trafficking, human organ trafficking, illegal drugs distribution. This paper introduces a novel approach that integrates humanized artificial intelligence (AI) with emotion-based threat prediction to enhance the detection and prevention of potential threats on the dark web. By analyzing emotional cues within dark web communications, the proposed system aims to identify and predict threats more effectively, thereby bolstering cybersecurity measures in this challenging domain.

Index Terms- humanized AI, Emotion Recognition, Threat Prediction, Dark Web, Affective Computing and Cybersecurity

I. INTRODUCTION

Anonymous transactions on the dark web are typically linked to data breaches, malware and other illegal activities. Classical Framework of Cybersecurity does not consider emotions and psychology. Humanized AI Humanized AI integrates emotion intelligence into AI systems and enables them to understand human emotions as human does by extracting sematic information conveyed in written language. In this paper, we explore how AI can be aware of emotions in respect to threats and determine when there is risk during dark web transactions.

II. LITERATURE REVIEW

At present, AI in the cybersecurity space is largely dominated by anomaly detection and sentiment analysis. Studies by Kumar et al. (2022) and Singh et al. 4353 4488) showed that the NLP could be applied to detect the extremist and dangerous language. Still, there is little evidence of affective computing that reads the subtler emotions. Recent works in transformer models such as BERT and Roberta have renewed the emotional detection on text data. In addition to these improvements, the proposed approach combines emotion recognition and threat intent classification into one model.

III. METHODOLOGY

The proposed system is divided into four main parts:
Gathered Data: Gathered data from dark web forums via ethical TOR relayed scrapping approach.
After the pre-processing (tokenize, stop word removal and lemmatization).
Emotion Recognition: It employed fine-tuned Roberta model on Go Emotions dataset to produce emotional embeddings.
Threat Prediction: These are forwards through emotional, linguistic and semantic features to a Blist Sub (subsequent) classifier layer.
The precision, recall and F1-score of the model were evaluated, In addition to its confusion matrix.

IV. RESULTS AND DISCUSSION

Humanized AI model yielded an accuracy of ~92%, better than the benchmarks like SVM and Random Forest. Emotional cues enhanced the identification of Time 1 signs of threat. Posts - involving aggression or

hostility were most likely linked to real cyber incidents which took place within 14 days. The experiments demonstrate that the emotion-aware model can contribute to situational awareness and improved prediction in cybersecurity.

V. CONCLUSION AND FUTURE SCOPE

In this paper, novelty, Humanized AI framework for emotion-based threat prognosis using affective computing and deep learning is presented. The introduction of sentiment makes the emotion based model more accurate and help us to understand better about input. Future work includes real-time applications, multimodal emotion assessment (text, audio and visual and address ethical challenges of emotionally intelligent monitoring systems.

REFERENCES

[1] Kumar, R., & Sharma, N. (2022). Threat detection using AI in dark web forums. IEEE Transactions on Cybernetics.

[2] Singh, A., & Kaur, P. (2023). Emotion recognition in online communication with BERT. Journal of Machine Learning Applications.