IJARSCT
ISSN (Online) 2581-9429
International Journal of Advanced Research in Science, Communication and Technology (IJARSCT)
Volume 2, Issue 1, June 2022
Impact Factor: 6.252

AI Based Self Learning Intelligent Information Leak Protection System for TI Companies using LSTM

K. Pazhanivel[1], E. Mounika[2], R. Shilpa[3], S. Sakthi[4]
Assistant Professor, Computer Science and Engineering[1]
Student, Computer Science and Engineering[2,3,4]
Anjalai Ammal Mahalingam Engineering College, Thiruvarur, India

Abstract: *Preventing the leak of sensitive information, also popularly known as data leak or data loss to an unauthorized recipient, is the primary goal of an organization's information security system. A data leak can occur through multiple channels. While it may not always be possible to prevent it entirely, measures can be taken to minimize the possibility of the occurrence. Like all other financial institutions, TI companies collect sensitive personal information of their customers for business purposes. This information is often categorized into three primary types; NPI, PII, and PI are the designated types in the descending order of sensitivity. The detection of sensitive documents and redaction of sensitive information is required if it is needed to be shared. Inspection of such digital documents to find any sensitive information is by far a human-driven process, and thus time-consuming and costly. An and costly. An intelligent and robust system is required where the content is analysed by state-of-the-art data mining, statistical and machine learning techniques from various data dimensions. An AI based self-learning Intelligent Information Leak Protection System using LSTM is proposed in the project that mines and extracts information and categorizes the document images, to SD or NSD, based on the presence of NPI and PII semantic signatures without any explicit rule configuration. The system is designed to be used proactively as an early warning system to tag the SD images while resting in the data store. It can also act as a real-time checkpoint for the information loss by the documents in transit or use. The proposed model prescribes an information loss protection mechanism using a binary classifier based on the state-of-the-art LSTM technique within the paradigm of Artificial Intelligence.*

Keywords: Critical Infrastructure Security, Network Security, Application Security, Cloud Security, Information Security, Disaster Recovery / Business Continuity Planning Storage Security, End-user Education, etc.

REFERENCES

[1] H. Alhindi, ``A framework for data loss prevention using document semantic signature," Ph.D. dissertation, Dept. Elect. Comput. Eng., Univ. Victoria, Victoria, BC, Canada, 2019.
[2] A. Guha and D. Samanta, ``Hybrid approach to document anomaly detection: An application to facilitate RPA in title insurance," Int. J. Autom. Comput., vol. 18, no. 1, pp. 55-72, Feb. 2021.
[3] Y. Lu, X. Huang, Y. Ma, and M. Ma, ``A weighted context graph model for fast data leak detection," in Proc. IEEE Int. Conf. Commun. (ICC), May 2018, pp. 1-6.
[4] K. Rishika and V. Damodaran, ``Data leakage detection: Challenges and prevention systems," Springer, Singapore, Tech. Rep., 2020.
[5] J. Zheng, ``Pattern matching for data leak prevention," US Patent 10 354 088, Jul. 16, 2019.R