Apoorva Joshi

San Jose, CA, USA
Email: appujo@gmail.com | **GitHub**: ajosh0504 | **LinkedIn**: apoorvajoshi95

Data Scientist with 4+ years of experience building and delivering production-level machine learning models to solve interesting problems in the Cybersecurity space.

Technical Skills

Programming Languages	Python, Java
Query Languages	SQL, Elasticsearch DSL
ML/DL Tools	Scikit-Learn, Pandas, NLTK, Gensim, Hugging Face, PyTorch, Keras
Data Visualization Tools	Kibana, Tableau, Matplotlib
Cloud Platforms	AWS, GCP
Miscellaneous	Elasticsearch, Docker, Git, FullStory, JIRA

Relevant Experience

Elastic San Jose, CA
Senior Data Scientist *March 2020-Present*

- Leading the delivery of user and entity behavior analytics (UEBA) features to the Elastic SIEM product
- Optimized a transformer model to detect malicious command lines, resulting in an $\approx 7\times$ speedup in inference time
- Developed a statistical framework to identify network beaconing, a common method used by malware to communicate with attacker-controlled servers
- Collaborated on a framework that uses supervised ML and anomaly detection to detect living-off-the-land attacks
- Analyzed customer telemetry data to understand the usage of Elastic's ML offerings, and communicated the findings to various stakeholders
- Designed a data sampling and anonymization tool to enable teams to safely use internal data for research, customer demos etc.

FireEyc Inc. San Jose, CA
Senior Research Scientist *July 2018-March 2020*

- Developed a Random Forest model for malicious URL detection, resulting in a 22% increase in the number of detections
- Used Transfer Learning and image clustering to build an algorithm for visually inspecting phishing campaigns

University of Iowa Iowa City, IA
Master's Thesis *August 2017-May 2018*

- Compared the performance of different ML models for classifying user posts mined from an online health community into sentiment categories
- Used a Bayesian approach on temporal user behavior trajectories to predict user churn from the community

Education

University of Iowa, *M.S. Computer Engineering* *May 2018*

VIT University, *B.E. Electronics and Communication Engineering* *May 2016*

Blogs, Presentations and Papers

- Blogs at Elastic can be found here.
- *ProblemChild in the Stack*. BSides Denver, March 2021 [Link]
- Joshi, A., (2019, October). *Using Lexical Features for Malicious URL Detection- A Machine Learning Approach*. Conference on Applied Machine Learning for Information Security (CAMLIS), Washington, DC, USA. [Link]
- Wang, X., Joshi, A., Zhao, K., Zhou, X., (2018, December). *Social Support and User Churn Prediction for Online Health Communities – A Trajectory-based Deep Learning Method*. 28th Workshop on Information Technology and Systems (WITS), California, USA.