

Project Ideas

Defensive

- WAF
- Anomaly detection in log & event files to determine whether there are irregularities that could be picked up on.
- Phishing detection
- Political deep-fake detection
- Email detection
- User behavior detection
- Malware detection
- AI driven IDS
- Endpoint monitoring
- Insider threat detection (rogue colleague)
- Password strength protection

Offensive

- Static code analysis
- Phishing generation
- User input sanitization bypass (payload creation)
- Blue-team evasion techniques

Short list:

1. WAF / Anomaly detection (L7)

Web app firewall that detects anomalies in web requests and blocks potential malicious activity. Model trained on known attack vectors and payloads. Also trained with good data when deployed.

2. Insider threat & phishing detection / email monitoring

Monitoring communication channels for phishing attacks (attachments, domains, social engineering in general)

3. Malware detection

Train a model to detect malware.

<https://www.vx-underground.org/malware.html>

<https://bazaar.abuse.ch/browse/>

4. AI driven IDS

Utilizing AI to monitor behavior and determine if an endpoint is compromised.

5. AI driven static code analysis

Research if it's possible for AI to create context that other SAST-tools miss (To reduce false-positives). Check what/where/when dangerous functions are used and if & how user-input+ could flow into them.

<https://snyk.io/blog/snyk-at-rsac-2021-ml-in-sast-distraction-or-disruption/>

<https://github.com/grassknotted/SAST-using-ML>

6. Phishing campaign generator

A red-teaming tool to create custom tailored phishing mails (maybe websites too), where a security researcher can give company/target specific parameters, trained on real emails.

7. Blue-team evasion techniques

Look into attack vectors like the one-pixel attack, to bypass detection techniques.

https://colab.research.google.com/github/hyperparticle/one-pixel-attack-keras/blob/master/1_one-pixel-attack-cifar10.ipynb#scrollTo=0YmVYTb6DLdK

Own project	Company project
6	6
3	1
1	3
2	2
4-5-7	