Data Lake & Brand Protection Predictive Model

Business Case:

Our customer is a leading service provider of digital brand services to the most prominent businesses globally. They have a huge responsibility of protecting their customers' brands and trademarks from various phishing channels. The client wanted a streamlined mechanism to gather, analyze, and report this data on a day-to-day basis.

Project Description:

- Gather and process data related to phishing of brands, trademark, domains, etc.
- Process the unstructured data and extract structure from it.
- Leverage an anti-phishing algorithm to assign proximity score and filter the probable phishing cases/URLs.
- Report final set of data to the agents (on a web portal) for further confirmation and take next course of action manually.
- Implement Big Data solution to handle high volume and unstructured data.

Platform Details:

- Private Data Center used
- Cloudera 5.8, Hadoop 2.x
- Solr, SonicMQ, Oracle

20 node cluster with:
- 100 GB RAM
- 12 core
- 5TB

Business Benefit:

- Automated anti-phishing service: Identifying the phishing content through multiple sources such as feeds, emails from the customers. The solution supports fraud prevention using a takedown process and monitors them after takedowns. It also protects the end-users by acting against phishing fraudsters who steal sensitive information for misuse.

- Internet Monitoring Service: Policing the internet to strengthen customer brand's credibility and value.

- Domain Name Monitoring: Monitoring domain registrations of exact matches, variants. Additionally, typos are identified, and reports are sent to the customers for further action.

- Image Match: Comparing the web images with customer images to identify any potential infringement or misuse of the customers' images related to the business.

- Analyst view for all customer-facing data. There is another eye watching the results, chances are very less for admitting mistakes in the end results and reduce the number of errors/erroneous data.

- Improved customer features (self-service) and reduced dial-in customer calls thereby achieving end-user satisfaction.

Statistics:

- Processed 800 Million records so far and landed in SOLR.
- On an average 250 sites are shut down due to this process across the globe, which is the number of average abuses.

Conceptual Design for ML:
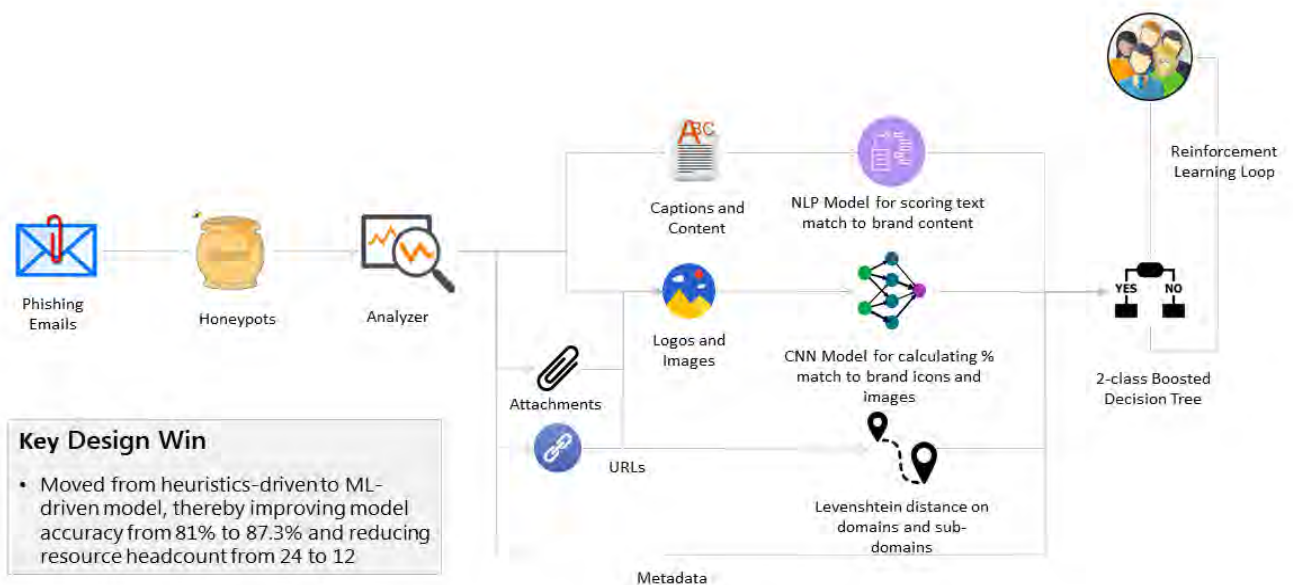
About Evoke

Evoke Technologies is an innovative information technology services firm offering world-class software solutions. We are a client-centric and relationship-based company with a focus to provide cost-effective software development and support solutions. We are SEI-CMMI Level 3 appraised, ISO 27001 and 9001 compliant IT enterprise that focuses on quality and delivering business value. Incorporated in the year 2003, Evoke Technologies has been actively helping global corporations to innovate and transform their businesses utilizing IT. Our global delivery and engagement models help us to offer affordable long-term software solutions that meet our client's IT challenges and add business value. We constantly innovate and implement new software methodologies to provide our clients with a superior experience.

EVOKE TECHNOLOGIES

📞 +1 (937) 660-4923

✉ sales@evoketechnologies.com

7106 Corporate Way,
Dayton, OH - 45459,
United States

www.evoketechnologies.com