PROPOSAL TUGAS AKHIR

NIM	Nama	Persentase Kontribusi
11320032	Ricky Ananda Pardomuan Sitorus	50%
11320043	Rut Ferwati Lumbantoruan	50%

Usulan Judul (Sementara) : Development Browser Extension for Detecting XSS Vulnerabilities based Website using Long Short Term Memory Attention (LSTM-Attention) Algorithm

Pembimbing : Togu Novriansyah Turnip, S.S.T., M.IM

Program Studi : D3 Teknologi Informasi

Jenis Tugas Akhir : Cyber Security

Mata kuliah yang pernah diambil terkait dengan Tugas Akhir ini:

- Keamanan Perangkat Lunak
- Teknologi Kecerdasan Buatan
- Pengujian Perangkat Lunak

A. Latar Belakang

Web adalah sistem yang digunakan untuk mengakses, mengunduh dan mengelola dokumen, teks, gambar, animasi, suara ataupun gabungan dari semua elemen yang menggunakan protokol HTTP [1]. Kumpulan dari halaman-halaman *web* statis maupun dinamis yang membentuk satu rangkaian yang saling terkait disebut dengan *website*. *Website* memiliki peran penting dalam perkembangan teknologi saat ini. *Website* sering digunakan sebagai wadah pencarian informasi, *online shopping*, kursus *online*, periklanan, dan *internet banking*. Penggunaan *website* kerat kaitannya dengan penggunaan internet karena *website* hanya bisa diakses melalui internet. Databoks mencatat bahwa ada sebanyak 4.95 miliar pengguna internet terhitung sampai Januari 2022. Data tersebut meningkat dari pengguna internet pada Januari 2021 sebanyak 4% dari 4.76 miliar [2]. Pengguna internet yang semakin tinggi memicu banyak tindakan kejahatan di dunia maya (*cybercrime*). *Cybercrime* adalah semua akses ilegal yang dilakukan menggunakan komputer untuk merusak, membahayakan, mencuri dan

menghancurkan data orang lain. Contoh *cybercrime* yaitu penyebaran virus atau *malware*, peretasan, pencucian uang, *phising*, dan *cyberstalking* [3].

Berdasarkan data *Open Web Application Security Project (OWASP),* ada 10 jenis serangan *cyber* yang paling sering menyerang keamanan *website*. Jenis serangan tersebut yaitu *SQL injection, Broken Authentication, XSS, XXE, Secure Misconfiguration, Sensitive Data Exposure, Insecure Deserialization, Using Components with Known Vulnerabilities,* dan *Insufficient Logging and Monitoring* [4]. XSS merupakan jenis serangan yang memungkinkan penyerang untuk mengeksekusi *script* di *browser* korban yang dapat membajak sesi pengguna, mencuri *cookie*, dan membocorkan informasi pribadi pengguna [4]. XSS terdiri dari 3 jenis, yaitu reflected-XSS (*non-persistent),* stored-XSS *(persistent),* dan DOM based-XSS. Reflected-XSS dilakukan dengan cara menyisipkan kode berbahaya ke URL yang dikunjungi oleh korban. Contohnya, penyerang akan membuat sebuah URL yang mengandung skrip berbahaya. Penyerang akan melakukan cara agar URL tidak mencurigakan, contohnya membuat URL menjadi lebih pendek. *User* akan mengakses URL tersebut dan berpikir bahwa URL tersebut merupakan URL yang normal. Tindakan yang dilakukan oleh *user* secara tidak langsung akan mengirimkan informasi penting *user* ke penyerang. *Developer* tidak tahu bahwa *website* yang dibangun telah dimasukkan skrip berbahaya oleh penyerang. Hal ini terjadi karena XSS digunakan untuk menyerang pengguna aplikasi, bukan aplikasi atau server [5]. Tahun 2020 ditemukan sebanyak 25% *web* yang memiliki kerentanan terhadap XSS [6]. Data tersebut meningkat sebanyak 1.5% dari tahun 2019. Peningkatan kerentanan XSS tersebut adalah bukti bahwa keamanan aplikasi *web* belum ditangani secara efektif sehingga diperlukan upaya untuk mengatasi kerentanan tersebut. Hal ini menjadi dasar yang digunakan dalam penelitian untuk meningkatkan keamanan aplikasi *web* terhadap kerentanan XSS.

Rathore dkk. [7] melakukan penelitian untuk mengklasifikasikan serangan XSS yang terjadi pada SNS (*Social Networking Service*) atau yang sering kita sebut dengan media sosial. Penelitian ini menggunakan perbandingan sepuluh algoritma seperti *Random Forest, ADTree, Random SubSpace, Decorate, AdaBoost.M1, JRip, Naive Bayes, Support Vector Machine, Logistic Regression, k-Nearest Neighbors.* Berdasarkan

penelitian yang telah dilakukan diperoleh hasil bahwa AdaBoost.M1 dan ADTree memberikan akurasi yang tinggi dalam mendeteksi XSS di lingkungan SNS. Namun, penelitian ini masih menggunakan *traditional machine learning* sehingga masih membutuhkan penelitian yang menerapkan metode lain dalam *machine learning* atau *deep learning* yang lebih akurat dalam mendeteksi XSS.

Lei Li dkk. [8] melakukan penelitian yang membandingkan empat algoritma yang digunakan dalam *deep learning*, yaitu RNN, GRU, LSTM, dan LSTM-Attention. Berdasarkan penelitian yang dilakukan, diperoleh hasil bahwa LSTM memiliki akurasi yang paling tinggi dalam mendeteksi XSS. Penelitian ini menambahkan mekanisme *attention* ke dalam model LSTM sehingga meningkatkan kinerja dalam pendeteksian XSS. Mekanisme *attention* ini menghitung probabilitas dari setiap kata-kata pada inputan dan memberi bobot untuk membedakan tingkat kepentingan informasi dari teks sehingga meningkatkan akurasi dari algoritma LSTM dalam mendeteksi XSS. Penelitian ini masih membutuhkan optimasi dalam melakukan pendeteksian dan pengklasifikasian XSS secara langsung pada jaringan.

Berbeda dengan penelitian Rathore dkk, penelitian ini akan membangun sebuah model menggunakan *deep learning* dengan algoritma LSTM-Attention (*Long Short Term Memory Attention*). LSTM melakukan proses urutan peristiwa melalui *cell memory*, memprediksi serta membandingkan data dengan informasi yang disimpan sebelumnya [9]. Algoritma LSTM digunakan untuk mengekstrak fitur secara otomatis dan melatih data *train* untuk mendeteksi XSS. Berdasarkan penelitian Lei Li dkk, mekanisme *attention* ditambahkan ke dalam model LSTM untuk meningkatkan kinerja dalam pendeteksian XSS dan akan digunakan pada penelitian ini. Model yang dibangun dalam penelitian ini akan menggunakan *dataset* Kaggle dan *dataset* SKYKAMI. Model yang sudah dibangun akan dihubungkan ke dalam sebuah *browser extension*. *Browser extension* digunakan untuk menambah fungsionalitas sehingga membantu pengguna untuk melakukan *browsing* dengan lebih mudah dan aman. Kegiatan *browsing* dapat dilakukan melalui *Firefox, Microsoft Edge*, *Opera*, *Google Chrome* dan lain-lain. Menurut data yang diperoleh dari Statcounter, *Google Chrome* merupakan *browser* yang paling sering digunakan di dunia. Data tersebut mencapai 66.3% dari jumlah

pengguna internet per September 2022 [10]. Tingginya penggunaan *Google Chrome* mengharuskannya untuk meningkatkan kualitas dari pelayanan yang tersedia, salah satunya adalah tingkat keamanan data pengguna. Salah satu kontribusi penelitian ini untuk meningkatkan keamanan data pengguna adalah dengan pengimplementasian *browser extension* yang akan dipasang pada *Google Chrome. Browser extension* yang diimplementasikan dalam penelitian ini akan membantu pengguna dalam mendeteksi URL yang mengandung XSS setiap melakukan *browsing.* Hal ini bertujuan untuk mengurangi kemungkinan pencurian data atau pengiriman data *cookie* secara ilegal. Sebuah *alert* akan ditampilkan untuk memberitahukan bahwa URL yang diakses mengandung XSS atau tidak.

B. Tujuan

Tujuan penelitian dalam pelaksanaan Tugas Akhir ini adalah membangun sebuah *browser extension* yang digunakan untuk mendeteksi XSS (XSS berjenis *reflected*) dengan menggunakan LSTM-Attention (*Long Short Term Memory Attention*).

C. Lingkup

Ruang lingkup dari pengerjaan tugas akhir ini yaitu:

1. Mengimplementasikan algoritma *deep learning,* yaitu algoritma LSTM-Attention dalam mendeteksi XSS
2. Model yang dibangun akan menggunakan *dataset* Kaggle dan *dataset* SKYKAMI
3. *Browser extension* akan dijalankan pada *Google Chrome*

D. Research Question(s)

Rumusan masalah dari penelitian ini sebagai berikut:

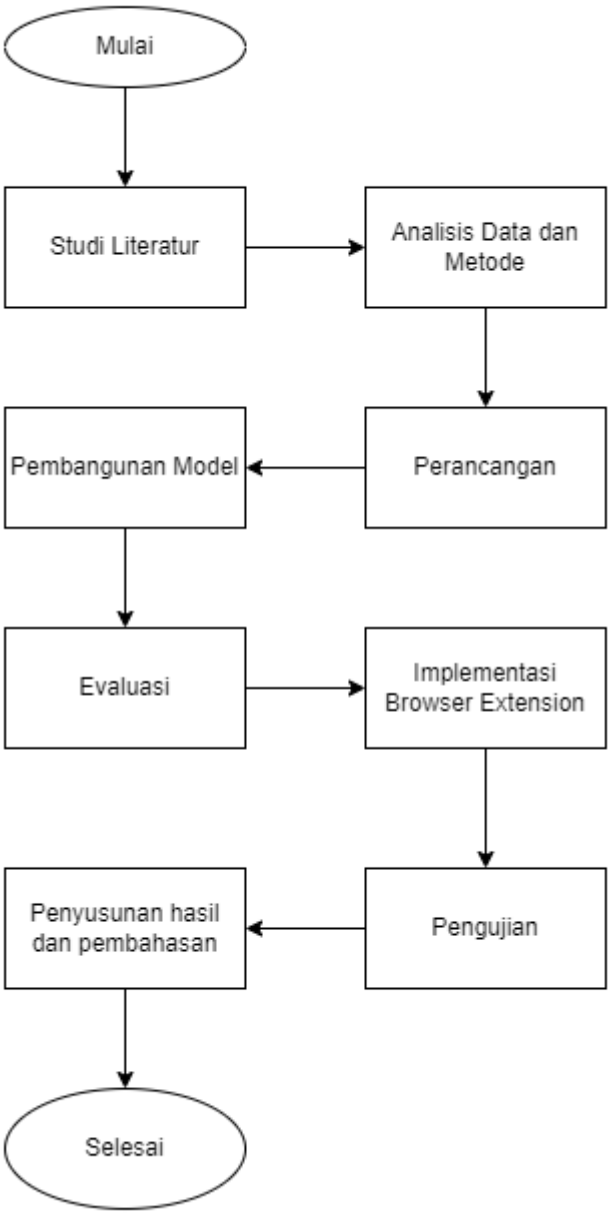
1. Bagaimana cara pembangunan *browser extension* dalam melakukan pendeteksian terhadap serangan XSS dengan menggunakan algoritma LSTM-Attention?
2. Bagaimana cara kerja LSTM-Attention dalam pendeteksian XSS pada *web*?
3. Bagaimana tingkat akurasi *browser extension* dalam mendeteksi serangan XSS?

E. Hasil yang Diharapkan

Hasil yang diharapkan pada Tugas Akhir ini adalah sebuah *browser extension* yang digunakan dalam pendeteksian URL yang mengandung XSS.

F. Metodologi Penelitian

Metodologi yang digunakan dalam menyelesaikan TA, digambarkan pada Gambar 1.



Gambar 1 Metodologi Penelitian

1. Studi Literatur

   Studi literatur merupakan tahapan awal dalam sebuah penelitian. Peneliti menyusun kerangka pemikiran atau suatu bahan bacaan sebagai pengantar dalam melakukan penelitian sesuai dengan topik TA. Hal ini bertujuan untuk menentukan daftar masalah yang akan diteliti sebagai landasan kegiatan. Bagian ini juga akan menjadi landasan untuk menggambarkan teknik dan metode serta algoritma yang akan digunakan dalam penelitian ini.

2. Analisis Data dan Metode

   Tahapan ini bertujuan untuk dapat memahami masalah yang sedang terjadi agar dapat menentukan metode yang tepat dalam memecahkan masalah. Pada tahapan ini juga akan dijelaskan terkait ketepatan dalam pemenuhan kebutuhan *user.*

3. Perancangan

   Tahapan ini bertujuan untuk memberikan gambaran dari *browser extension* yang akan dibangun. Pada tahapan ini akan dilakukan perancangan metode pembangunan model dan bahasa pemrograman yang digunakan untuk melakukan implementasi.

4. Pembangunan Model

   Tahapan ini bertujuan untuk membangun model yang akan digunakan dalam pengujian XSS dan digunakan dalam melakukan implementasi dari *browser extension.* Model tersebut akan dibangun menggunakan algoritma LSTM-Attention (*Long Short Term Memory Attention*)

5. Evaluasi

   Tahap evaluasi merupakan tahap dimana peneliti akan mengamati dan menguji model yang sudah dibangun. Model akan dievaluasi dengan menggunakan metrik. Hasil pengujian akan menemukan model dengan akurasi tertinggi. Model dengan akurasi tertinggi akan digunakan dalam implementasi *browser extension.*

6. Implementasi *Browser Extension*

   Tahapan ini merupakan realisasi dari pembangunan model dan *browser extension.* Semua perancangan yang sudah dijelaskan sebelumnya akan diimplementasikan pada tahapan ini. *Browser extension* akan dibangun menggunakan bahasa pemrograman *JavaScript* dan dihubungkan dengan model yang sudah dibangun sebelumnya.

7. Pengujian

   Pengujian dilakukan untuk memastikan elemen dan komponen dari *browser extension* yang sudah dibangun berfungsi sebagaimana yang diharapkan. Tahapan pengujian akan dilakukan dengan cara memasukkan URL yang mengandung XSS. Jika *browser extension* bisa mendeteksi XSS dan menampilkan *alert* maka tujuan dari penelitian ini tercapai.

8. Penyusunan Hasil dan Pembahasan

   Pada tahap ini akan dilakukan penyusunan hasil mulai dari perencanaan hingga ke hasil pengujian *browser extension*. Bagian ini akan menjelaskan kesimpulan dari pembangunan aplikasi yang dilakukan berdasarkan tujuan dan rumusan masalah yang telah diteliti.

G. Risiko

Risiko yang mungkin terjadi dalam pengerjaan tugas akhir ini adalah ketidaksesuaian hasil yang diperoleh. Implementasi yang dilakukan tidak memberikan hasil yang memuaskan dan tidak sesuai dengan tujuan.

H. Rencana Kerja

Jadwal rencana kerja dalam pengerjaan TA dapat dilihat pada Tabel 1.

Tabel 1 Rencana Kerja TA

Kegiatan	Minggu																	
	1	2	3	4	5	6	7	8	9	10	11	12	13	14	15	16	17	18
Konsultasi judul dengan dosen pembimbing																		
Requirement gathering dan studi literatur																		
Analisis data dan metode																		
Perancangan																		
Review dan revisi proposal dari pembimbing																		
UTS																		
Seminar Proposal																		
Revisi proposal sesuai dengan seminar proposal																		
Pembangunan Model																		
Pengerjaan Laporan TA Bab I dan Bab II																		
Evaluasi, Pengerjaan Laporan Bab III																		
UAS dan Seminar TA 1																		
Perbaikan dan melakukan finalisasi Laporan TA 1																		
Submit Artefak TA 1																		

Jadwal rencana kerja dalam pengerjaan TA 2 dapat dilihat pada Tabel 2.

Tabel 2 Rencana Kerja TA 2

Kegiatan	Minggu																
	1	2	3	4	5	6	7	8	9	10	11	12	13	14	15	16	17
Konsultasi dengan dosen pembimbing																	
Memperbaiki laporan TA Bab I - III berdasarkan hasil seminar																	
Melakukan eksplorasi terhadap tools yang digunakan																	
Implementasi																	
UTS																	
Implementasi																	
Pengujian																	
Pengerjaan Laporan TA Bab IV dan Bab V																	
Finalisasi laporan TA																	
UAS dan Seminar TA 2																	
Revisi TA 2 (Jika ada)																	

I. Referensi

[1] M. R. Arief, *Program web dinamis menggunakan php dan mysql*. Yogyakarta: Gratia K, 2011.

[2] "Pengguna internet di dunia capai 4,95 miliar orang per januari 2022." https://databoks.katadata.co.id/datapublish/2022/02/07/pengguna-internet-di-dunia-capai-495-miliar-orang-per-januari-2022 (accessed Oct. 05, 2022).

[3] W. A. Al-Khater, S. Al-Maadeed, A. A. Ahmed, A. S. Sadiq, and M. K. Khan, "Comprehensive review of cybercrime detection techniques," *IEEE Access*, vol. 8, pp. 137293–137311, 2020, doi: 10.1109/ACCESS.2020.3011259.

[4] OWASP, "OWASP top 10 - 2017 the ten most critical web application security risks," *OWASP Found.*, pp. 1–24, 2017, [Online]. Available: https://owasp.org/www-project-top-ten/2017/Top_10

[5] S. Suroto and A. Asman, "Ancaman terhadap keamanan informasi oleh serangan cross-site scripting (XSS) dan metode pencegahannya," *Zo. Komput.*, vol. 11, no. 1, pp. 11–19, 2021, [Online]. Available: http://www.hackers.com?yid=

[6] A. Report, "The invicti appsec indicator spring 2021 edition : acunetix web vulnerability report introducing the invicti appsec indicator," p. 45, 2021, [Online]. Available: https://www.acunetix.com/wp-content/uploads/2021/04/Invicti-AppSec-Indicator-Spring-2021-Edition-Acunetix-Web-Vulnerability-Report.pdf

[7] S. Rathore, P. K. Sharma, and J. H. Park, "XSSClassifier: an efficient xss attack detection approach based on machine learning classifier on snss," *J. Inf. Process. Syst.*, vol. 13, no. 4, pp. 1014–1028, 2017, doi: 10.3745/JIPS.03.0079.

[8] L. Lei, M. Chen, C. He, and D. Li, "XSS detection technology based on LSTM-attention," *2020 5th Int. Conf. Control. Robot. Cybern. CRC 2020*, pp. 175–180, 2020, doi: 10.1109/CRC51253.2020.9253484.

[9] H. Weytjens and J. De Weerdt, "Process outcome prediction: CNN vs. LSTM (with

attention),” *Lect. Notes Bus. Inf. Process.*, vol. 397, pp. 321–333, 2020, doi: 10.1007/978-3-030-66498-5_24.

[10] “Desktop, tablet & console browser market share worldwide | statcounter global stats.” https://gs.statcounter.com/browser-market-share/desktop-tablet-console/worldwide/#monthly-202109-202209 (accessed Oct. 26, 2022).