# Attention for detection SQL-injection queries

Mahmoud M. Saad[1] and Ahmed C. Kaseb[2]

[1] *Cairo University, Computer Department, Cairo, Egypt*
[2] *Cairo University, Computer Department, Cairo, Egypt*

Reception date of the manuscript: 12/12/2022
Acceptance date of the manuscript: 20/1/2023
Publication date: 23/1/2023

**Abstract**— SQL injection detection systems are considered the most valuable part of any network systems nowadays. Any SQL injection attack can harm every sensitive data stored over databases. SQL queries are used to interact with the database management system and retrieve data from the database. SQL query are optimized to be suitable and interactive with database system structure. Any Attack with SQL injection can access the database with the 4 known operation(CRUD) create, read, update/modify, delete data from server. for that it causes huge damage when the attack penetrate the system. Artificial intelligence and machine learning techniques have been tested and used to control SQL injection attacks, showing promising results. The main contribution of this paper to provide a deep neural network Attention based SQL injection detection model. As per literal review [1] all methods review were the popular architecture known in the field of deep learning as RNN, LSTM, CNN and their variation. Some work include ensemble learning with gathering more than one architecture during run time and decide based on highest votes but None include attention. Attention mechanism showed astonishing performance [2] and SOTA performance over many popular NLP tasks. based on that, in this paper we will consider implement SQL detection system with attention mechanism included in the architecture to give the model more ability to capture the context between important keywords during processing the SQL query and understand the patterns that can cause harmful SQL operations. our attention model using during this paper is BERT(Bidirectional Encoder Representations from Transformers) the BERT transformer is working with the input sequence in parallel and to keep the sense of sequence order, positional encoding embedding is used while embedding the word to vectors before feeding to model network. dropouts and activation layers implemented in the model to avoid fitting and gradient problems. As this problem is classification problem the cross entropy loss used as cost function to penalize the model when misclassification. for evaluation the models and measure the accuracy, confusion matrix was developed to draw out the strength and weakness of classification of the models through the False Positive and False Negative counts. Using attention through Transformers improved the accuracy (with detection accuracy of) of SQL injection detection systems while giving the models more understanding about the nature of injection rather than only learn if the query is injection or normal query as we got from the traditional classification methods

**Keywords**— First word or key phrase, second word or key phrase, third word or key phrase. (Place between three and six key words or phrases separated by a comma, which represent the theme of your work)

## I. Introduction

## II. Related Works

## III. Model design

Our Model used here is BERT(pre-trained of Deep Bidirectional Transformers). the motivation goes as we saw in the related work that no contribution before used attention (reference to attention is all what you need). Using attention is NLP application showed many SOTA results over many NLP popular applications. Transformers are build over the attention technique where we have encoder and decoder part as figure(x). Due do existence of huge corpus for Natural language processing. Transformers are trained over huge amount of data to get the best representation and understanding for the text. after that it's used as pretrained model to be fine-tuning over specific application dataset. the transformers showed great capabilities for capturing sentences context and correlations between words. from that effort had been done to embedded BERT to the research area of SQL injection.

the normal sequence of NLP application goes as following: (a) we do data cleaning and data processing to remove any redundant text and fix any issues exist in the corpus. this step also include truncate the sentences to have a fix length. (b) Embedding the text, In this step the text convert to a series of numbers so to be feed to the neural network module(RNN, LSTM) throught different embedding techniques (c) model implementation through the modern deep learning framework (Pytorch, tensorflow, caffee), here an explicit implementation for the model is done and the model inference flow step by step. associated with it the training loop where

Contact data: Mahmoud M. Saad, mahmoud.saadeldin@gmail.com

we define the stochastic gradient descent optimizer that do the most important part of training which is the gradients calculating throught the backward propagation step. (d) training analysis step in which we trigger any issue during training e.g. fitting problem: over-fitting and under-fitting problems, Gradient problems: vanshing and exploding gradients.

Bert has different terminology where the pre-trained model has different way to deal with the input sequence. (a) associated with the pretrained mode different tokinezers to split the sentence to separate words. After that those words converted to IDs as defined at the pretrained model. (b) the Bert model process the sequence in parallel way where we feed the sequence all at once in opposite to LSTM where we can't process word until we process the previous word. Bert does positinal encoding to give attention to the order of the sequence while parallel processing the sequence.

## IV. Results and Discussion

## V. Conclusions

## References