

**1<sup>ST</sup> INTERNATIONAL CONFERENCE ON  
INFORMATION AND COMMUNICATION TECHNOLOGY FOR DEVELOPMENT (ICICTD 2022)**

**July 29 - 30, 2022**

**Institute of Information and Communication Technology (IICT)  
Khulna University of Engineering & Technology (KUET)  
Khulna-9203, Bangladesh**

# **Effectiveness of Transformer Models on IoT Security Detection in StackOverflow Discussions**

## **Authors:**

- 1. Nibir Chandra Mandal**
- 2. G. M. Shahariar**
- 3. Md. Tanvir Rouf Shawon**

## **Presented by -**

**Md. Tanvir Rouf Shawon  
Ahsanullah University of Science  
and Technology  
Bangladesh**



# Contents of Presentation

- Introduction
- Problem Statement
- Related Works
- Objectives
- Methodology
- Results and Discussions
- Conclusions
- Future Works
- References

# Introduction

- ❑ IoT (Internet of Things) - a widespread connection of intelligent devices.
- ❑ A deliberate intrusion into the device might cause significant harm to the users.
- ❑ Discussion about the security of IoT devices in developer communities like Stack Overflow (SO).
- ❑ Extensive experiment for determining whether a discussion is an IoT security discussion or not.
- ❑ Proposed a domain-specific dataset focused solely on IoT security issues - "IoT Security Dataset".
- ❑ Experimented with transformer models like BERT, RoBERTa, XLNet and BERTOverflow in 3 of our different experiments.

# Problem Statement

- ❑ **A comprehensive analysis of determining whether an aspect of a discussion in Stack Overflow is an IoT security aspect or not.**
- ❑ **Assist developers in learning more about security issues in the IoT space, where there is a lot of discussion about various aspects.**
- ❑ **Vendors may increase the security of their products.**

# Related Works

- ❑ Uddin et al. [1] investigated at nearly 53,000 IoT-related posts on StackOverflow and used topic modeling to figure out what people were talking about.
- ❑ Aly et al. [2] addressed at questions of IoT and Industry 4.0 on StackOverflow. They also utilized topic modeling to identify the themes mentioned in the investigated questions, similar to the work given by Uddin et al. [1].

# Related Works

- ❑ Uddin and Khomh [3] introduced OPINER, a method for mining API-related opinions and providing users with a rapid summary of the benefits and drawbacks of APIs when deciding which API to employ.
- ❑ Recent studies [4][5] explored the connection between API usage and Stack Overflow discussions. Both research discovered a relationship between API class use and the number of Stack Overflow questions answered.

# Related Works

- ❑ Uddin and Khomh [3] utilized their constructed benchmark dataset named "OPINER" to carry out the study and noticed that developers frequently provided opinions about vastly different API aspects in those discussions, which was the first step towards filling the gap of investigating the susceptibility and influence of sentiments and API aspects.

# Related Works

- ❑ Based on the positive and negative results emphasized in earlier attempts to automatically mine API opinions, as well as the seminal work in this field by Uddin and Khomh [6], Lin et al. [7] introduced a new approach called Pattern-based Opinion Miner (POME), which utilizes linguistic patterns preserved in StackOverflow sentences referring to APIs to classify whether a sentence is referring to a specific API aspect and has a positive or negative polarity.



# Objectives

- ❑ The objectives of this paper are as follows:
  - ❑ To create a domain-specific dataset called "IoT Security Dataset" that focuses on security aspects of IoT related textual developer discussions.
  - ❑ To employ multiple transformer models to automatically detect security discussions. Through rigorous investigations, we found that IoT security discussions are different and more complex than traditional security discussions.
  - ❑ To demonstrate a considerable performance loss (up to 44%) of transformer models on cross-domain datasets while transferring knowledge from a general-purpose dataset to support our claim.

# Methodology

## Architecture of the security aspects classifier

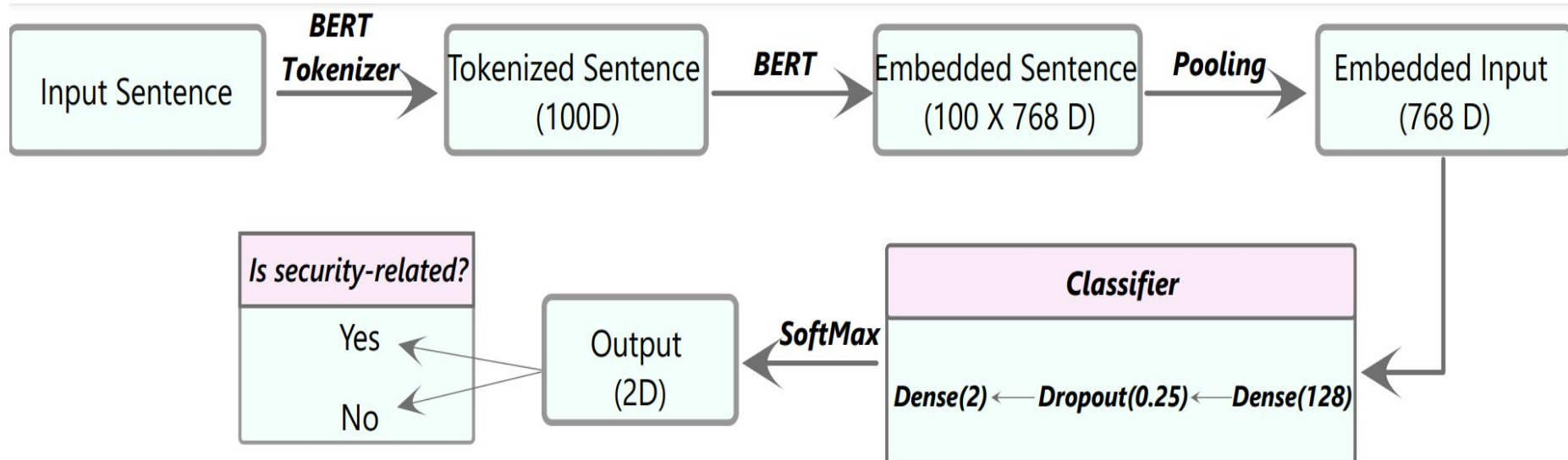


Fig: Aspects Classification Process

# Methodology Cont'd

## ❑ Experiments

- ❑ Transformer performance on

  - ❖ Opiner Dataset

  - ❖ IoT dataset

  - ❖ Combined dataset

## ❑ Evaluation

- ❑ Precision

- ❑ Recall

- ❑ F1-Score

- ❑ Area Under the Curve (AUC)

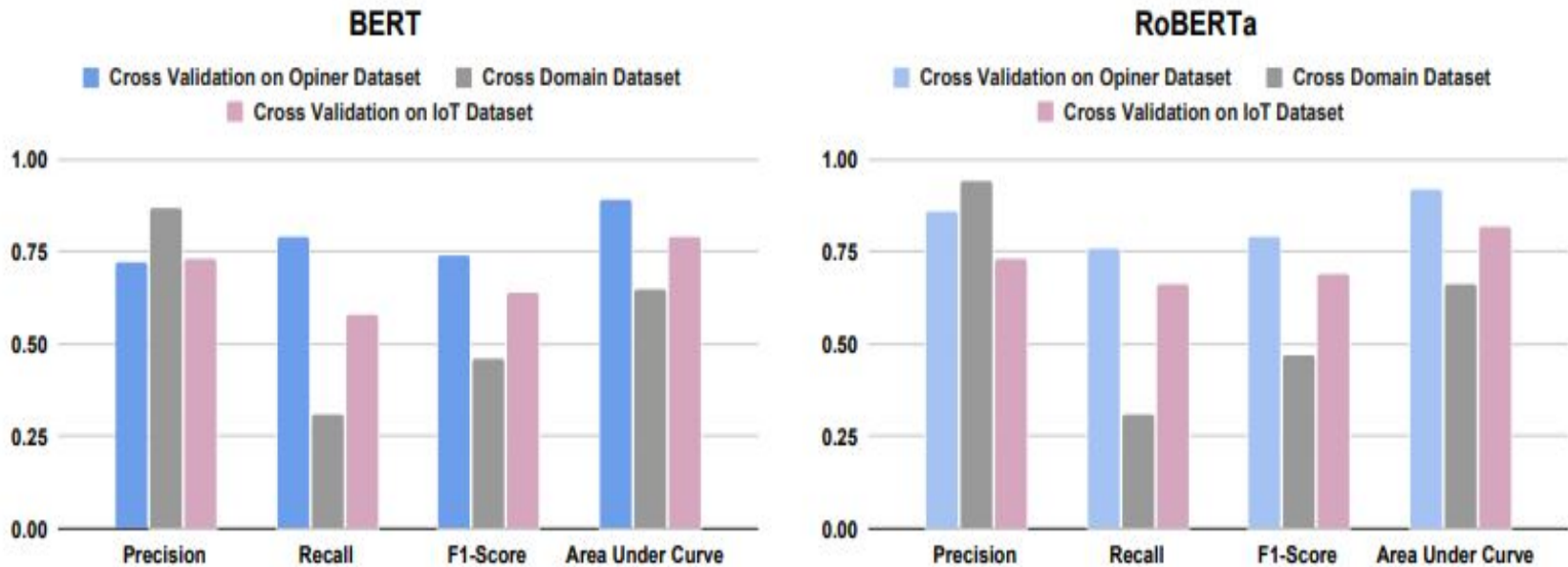
# Results and Discussions

- ❑ RoBERTa achieves highest 0.79 F1-Score
- ❑ Proposed models improve the benchmark i.e. Logits

Type	Model	Precision	Recall	F1-Score	AUC
Proposed Models	BERT	0.72	0.79	0.74	0.89
	RoBERTa	<b>0.86</b>	0.76	<b>0.79</b>	<b>0.92</b>
	XLNet	0.72	<b>0.84</b>	0.77	0.91
	BERTOverflow	0.79	0.76	0.76	0.88
Previous Work	Logits [19]	0.77	0.57	0.60	0.69

Fig: Performance of 4 different transformer model on Opiner dataset

# Results and Discussions – Cont'd



**Fig: A comparison of performance among all the 3 experiments for BERT (left) and RoBERTa (Right)**

# Results and Discussions – Cont'd

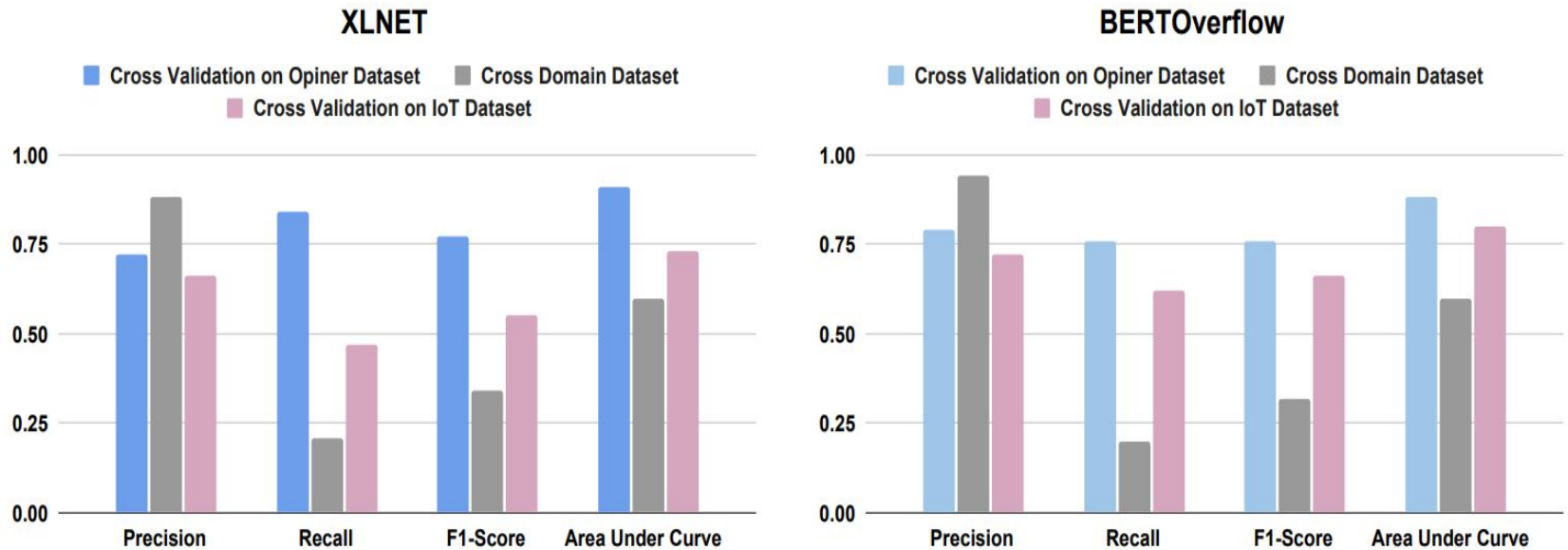
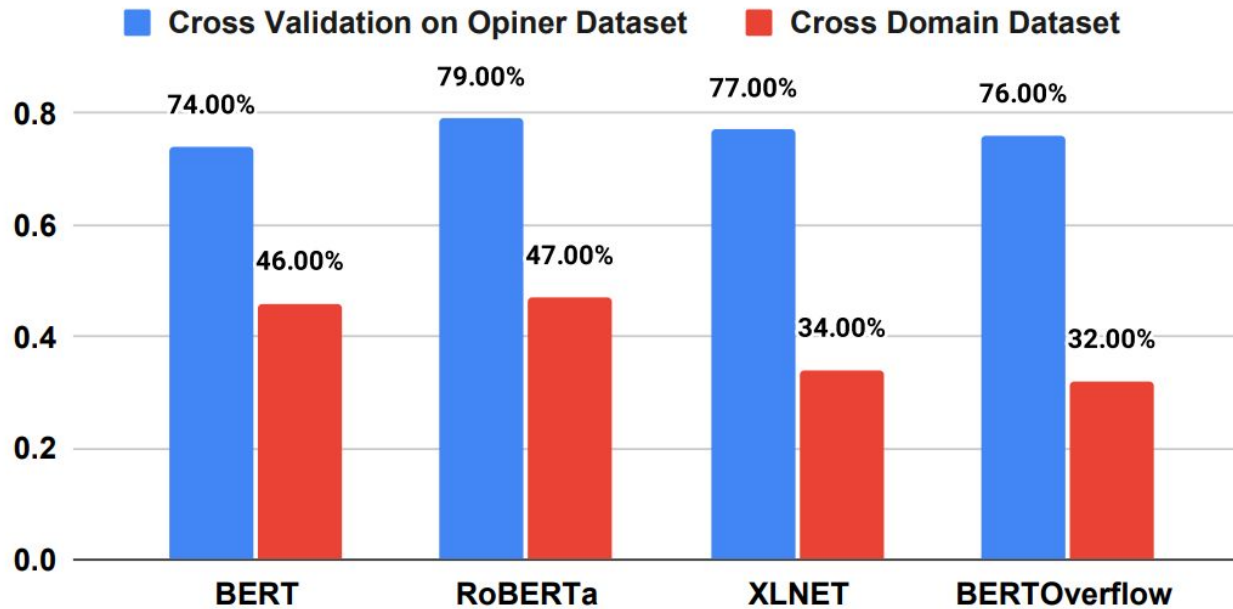


Fig: A comparison of performance among all the 3 experiments for XLNet (left) and BERTOverflow (Right)

# Results and Discussions – Cont'd

**Comparison of F1-Score between Cross Validation on Opiner Dataset and Cross Domain Dataset**



**Fig: A comparison of F1-Score between the experiments on Cross-validation on Opiner dataset and Cross-Domain dataset**

# Results and Discussions – Cont'd

- ❑ **IoT Security is different from general security discussion**
  - ❑ **Transformer performances drop in IoT domain**
  - ❑ **Performance can improve if IoT domain knowledge is provided**
- ❑ **IoT security aspects are more complex, sparse, and implicit**
  - ❑ **Adaptation of IoT knowledge in Transformer doesn't guarantee highest performance**



# Conclusions

- ❑ **Attempt to create a one-of-a-kind dataset.**
- ❑ **Comparison between our IoT security dataset and the benchmark dataset.**
- ❑ **Generalization is not really the best method for identifying security discussions.**
- ❑ **Domain-specific knowledge transfer via various transformer models might be a superior alternative to security aspect detection.**

# Future Works

- ❑ Incorporate other transfer learning models to improve our performance.
- ❑ Increasing the number of samples in the dataset is another effort we may undertake in the future to enhance the outcome.

# References

- [1] Uddin, Gias, Fatima Sabir, Yann-Gaël Guéhéneuc, Omar Alam, and Foutse Khomh. "An empirical study of IoT topics in IoT developer discussions on StackOverflow." *Empirical Software Engineering* 26, no. 6 (2021): 1-45.
- [2] M. Aly, F. Khomh, and S. Yacout. "What do practitioners discuss about IoT and industry 4.0 related technologies? characterization and identification of IoT and industry 4.0 categories in Stack Overflow discussions." *Internet of Things*, 14:100364, 2021.
- [3] Gias Uddin and Foutse Khomh. "Mining API aspects in API reviews." *Technical report*, 2017.
- [4] D. Kavalier, D. Posnett, C. Gibler, H. Chen, P. Devanbu, and V. Filkov. "Using and asking: APIs used in the Android market and asked about in StackOverflow." *In Proceedings of the International Conference on Social Informatics*, pages 405–418, 2013.

# References

- [5] C. Parnin, C. Treude, L. Grammel, and M.-A. Storey. "Crowd documentation: Exploring the coverage and dynamics of api discussions on stack overflow." *Technical report, Technical Report GIT-CS-12-05*, Georgia Tech, 2012.
- [6] Gias Uddin and Foutse Khomh. Opiner: an opinion search and Gias Uddin and Foutse Khomh. "Opiner: an opinion search and summarization engine for apis." *In Proceedings of the 32nd IEEE/ACM International Conference on Automated Software Engineering (ASE 2017)*, pages 978–983. IEEE Computer Society, 2017.
- [7] Lin, Bin, Fiorella Zampetti, Gabriele Bavota, Massimiliano Di Penta, and Michele Lanza. "Pattern-based mining of opinions in Q&A websites." *In 2019 IEEE/ACM 41st International Conference on Software Engineering (ICSE)*, pp. 548-559. IEEE, 2019.

# Thank You

## Any Question?

