UTSA Computer Science

Using Deep Learning to Bridge the Semantic Gap between Natural Language and Code: An Automated Approach to Improve Privacy Policy Misalignment Scalability

Chris Crabtree, Rocky Slavin, Xiaoyin Wang, and Jianwei Niu

Android Privacy Policy Problems

- Semantic disconnect between natural language policies and source code
- Varying styles of policy representation
- Varying coding habits
- Violations occur when data gathered by an app contradicts what is stated in a privacy policy

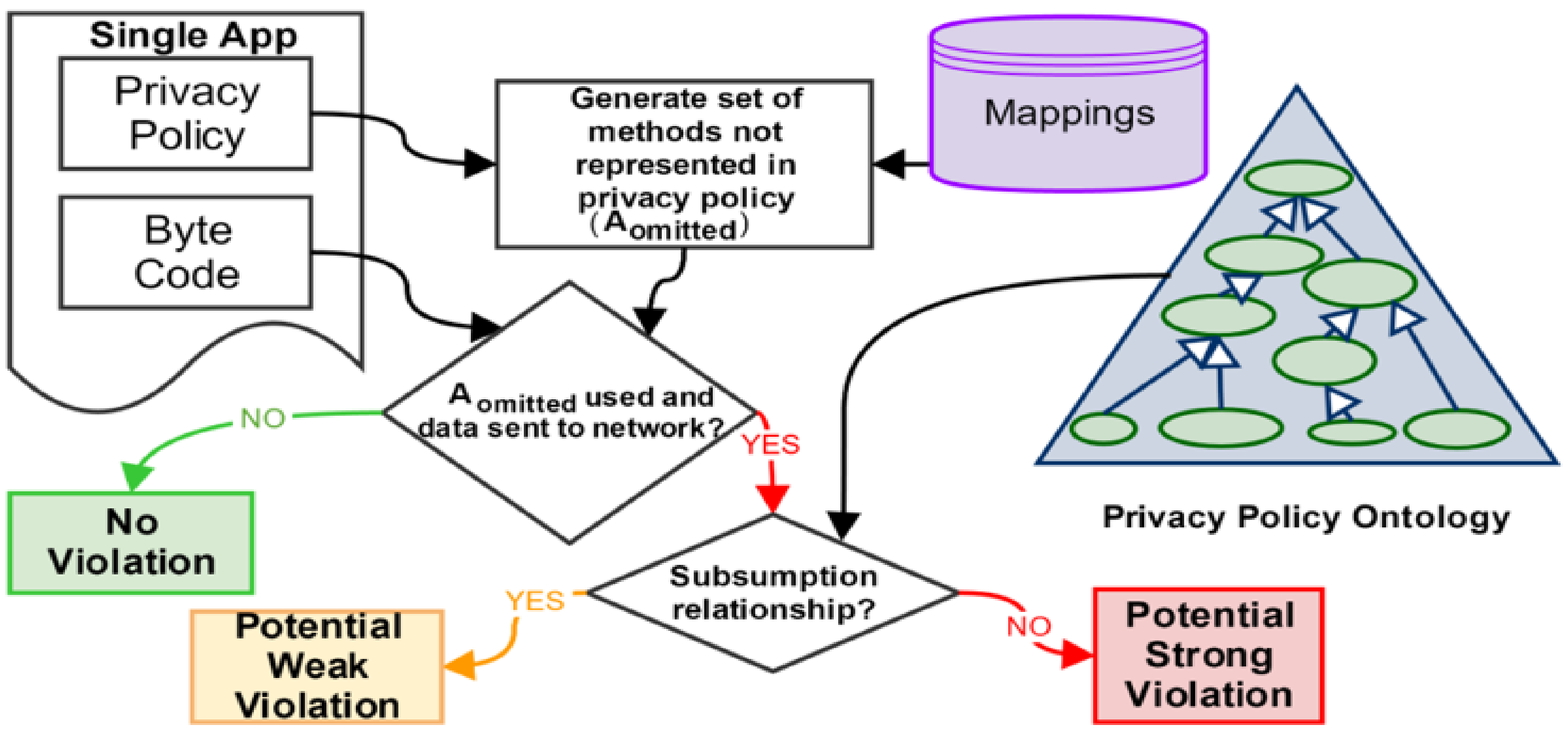
Single App
Privacy Policy
Byte Code
Generate set of methods not represented in privacy policy (A omitted)
Mappings
A omitted used and data sent to network?
NO
YES
No Violation
Subsumption relationship?
YES
NO
Potential Weak Violation
Potential Strong Violation
Privacy Policy Ontology

Figure 1. Violation detection framework

Motivation

- Previous work has been done on creating a framework for violations detection (Figure 1)
- A critical component is mapping data types described in privacy policies to methods in the Android Application Programming Interface (API)

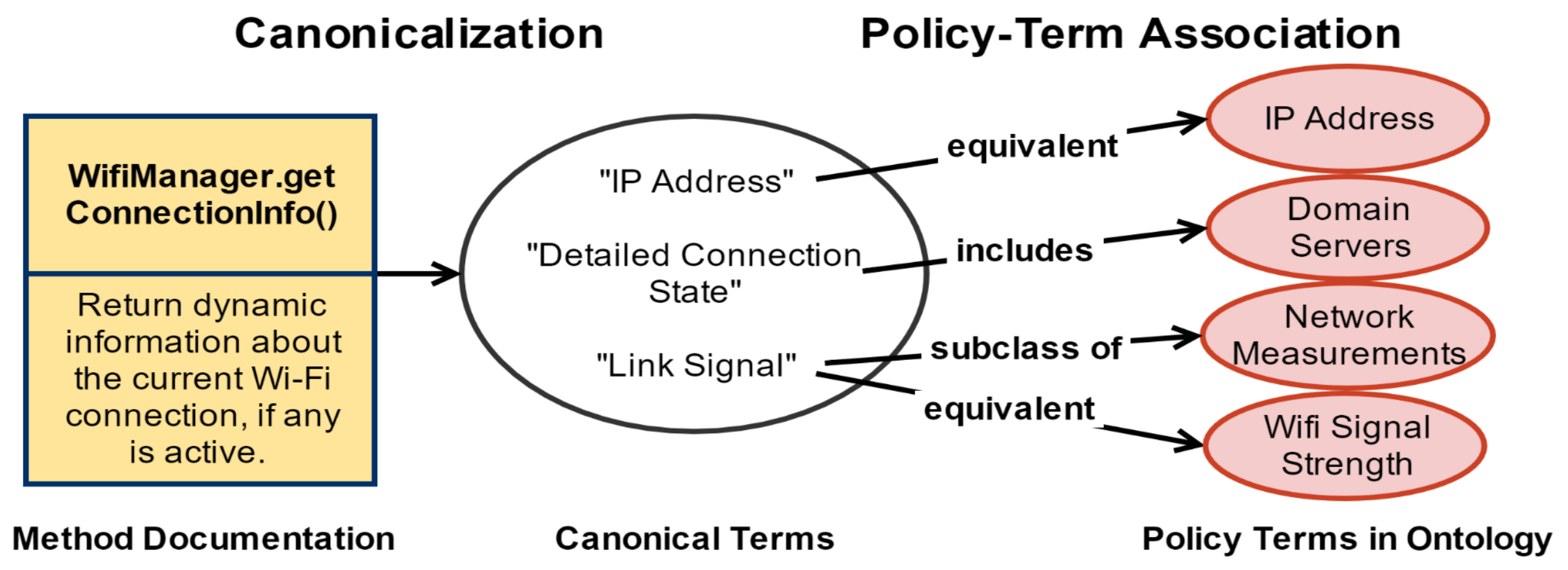

Figure 2. Previous mapping model. All mappings created manually.

Old Model (Figure 2)

- Intermediate ‘Canonical Terms’ were constructed.
- Android API methods were annotated by Canonical Terms, then Canonical Terms were mapped to privacy policy data types
- ***Each step was done manually***

Automated Natural Language Mapping Approach

- Use a deep leaning model to learn semantic associations.
- Utilize transfer learning to overcome small dataset
- A model trained on a similar task with a large amount of data can ‘transfer’ what it knows to a new task.
- The Bidirectional Encoder Representations for Transformers’ (BERT) model is a SOTA language model designed to utilize transfer learning (Figure 3).

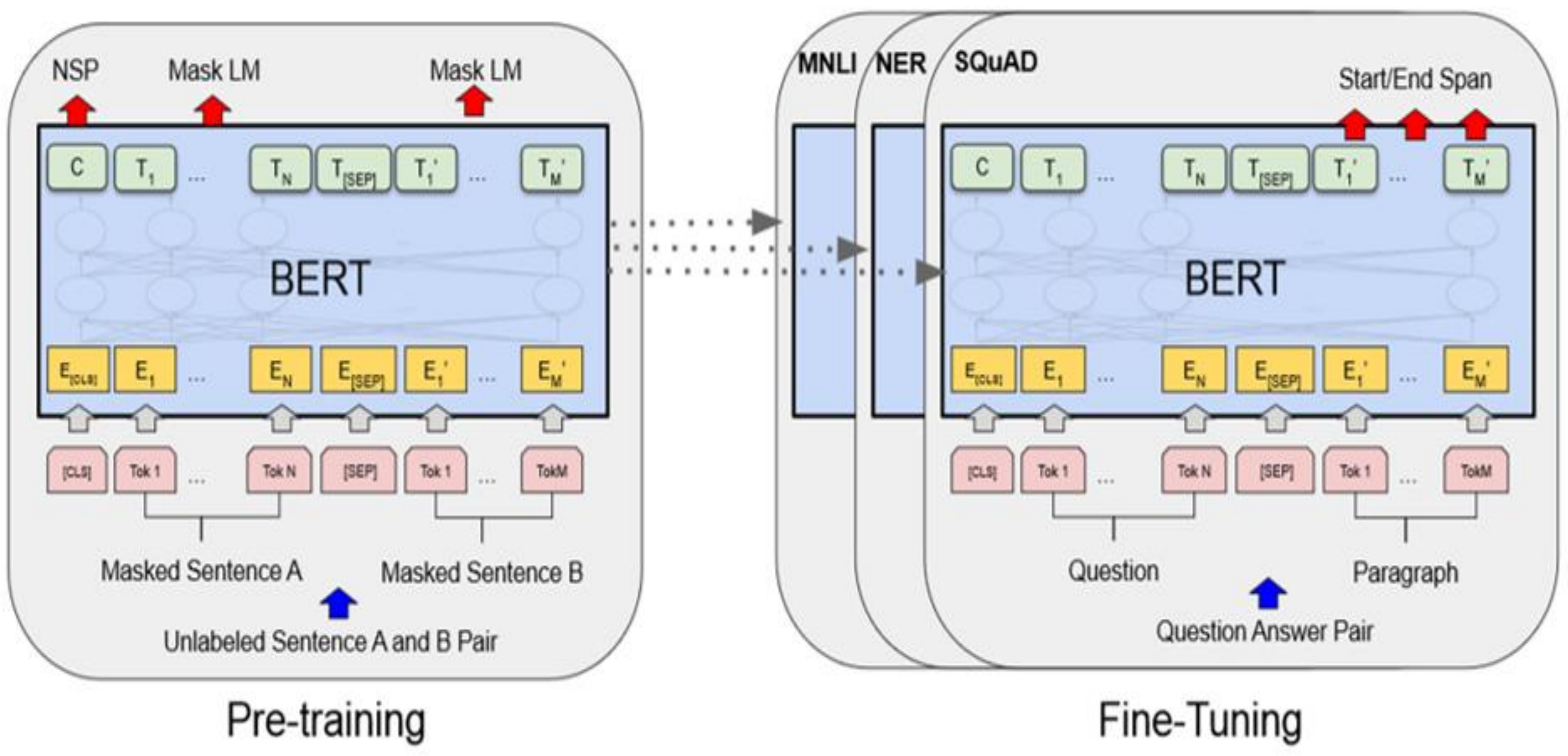

Figure 3. Transfer Learning with BERT model.

Source: BERT: Pre-training of Deep Bidirectional Transformers for Language Understanding (Devlin et al. 2018)

New Model (Figure 4)

- We jointly trained a feed forward multilabel classifier from the output of the BERT language model to deal with having few training examples. (Figure 4)
- To augment our training examples further, we split each method description into its sentences and added a sentence constructed from the method title

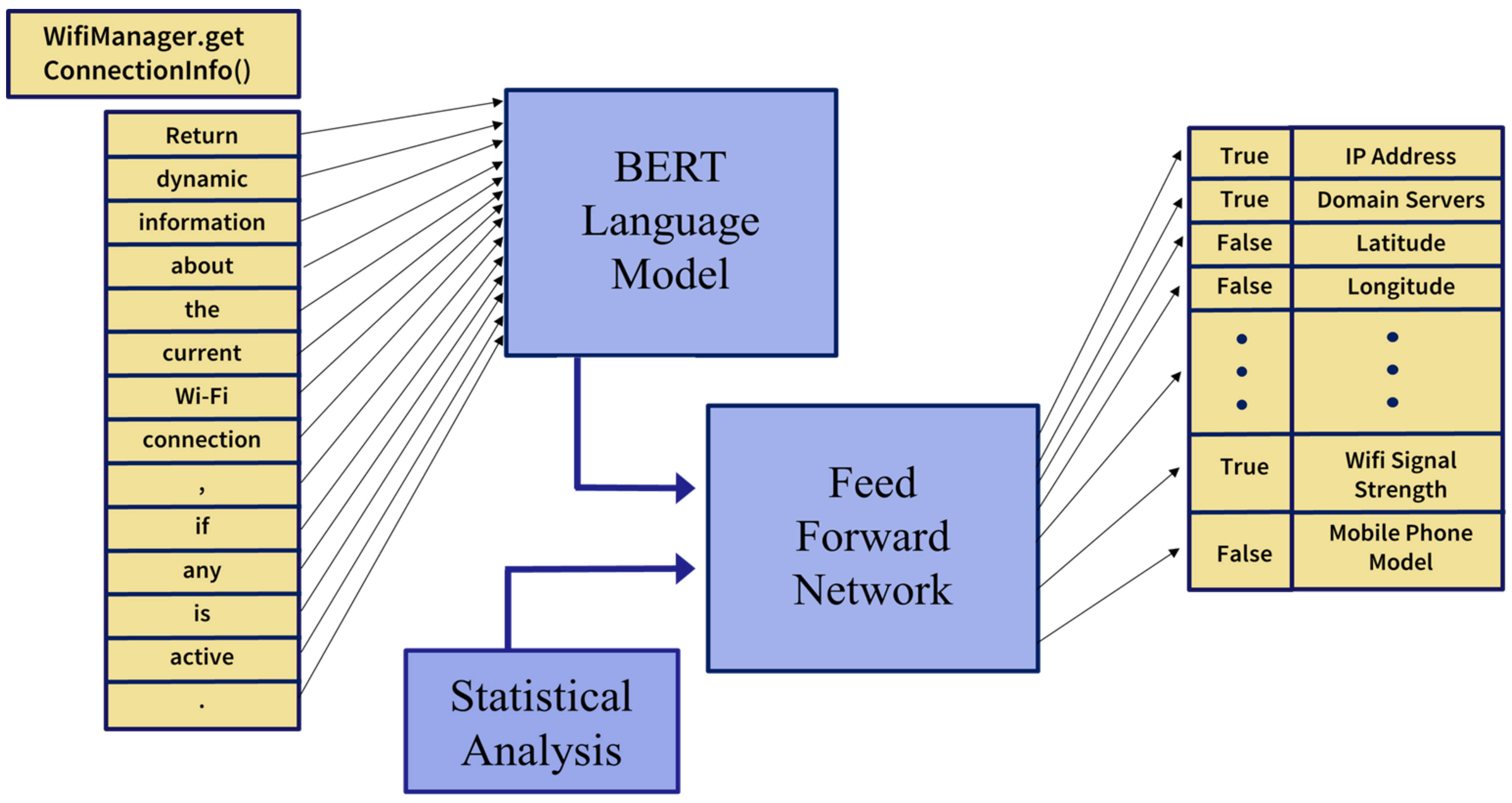

Figure 4. Deep learning model for privacy policy to documentation mapping.

Impact

- Automation means privacy policy violation detection will be robust to updates in the Android API
- Dramatically reduces resource usage
- Broader potential coverage for violation detection
- Creates framework for expansion of privacy policy data types
- Increases privacy compliance

NSF
The 4th NSF Secure and Trustworthy Cyberspace Principal Investigator Meeting (2019 SaTC PI Meeting)
October 28-29, 2019 | Alexandria, Virginia