# Real-Time Privacy Risk Detector for Android Apps

Palanivel Rajan C[1*], Nagalakshmi B[2], and Roshini S[3]

[1,2,3]Department of Information Technology, Sri Krishna College of Technology, Coimbatore, India
[1*]pvr2684@gmail.com
[2]naagalakshmi18.08@gmail.com
[3]roshinivshan@gmail.com

**Abstract.** Android applications are regularly involved with sensitive user information in the contemporary mobile ecosystem and are often regulated by privacy policies that are too long, vague, and incomprehensible to the end-user. Such ambiguity may result in the accidental consent, unauthorized data gathering, third-party surveillance, and the breach of the regulatory requirements like GDPR and CCPA. As a counter to these privacy issues, this paper has presented a prototype of a Real-Time Privacy Risk Detector of Android Apps which is an auto- mated system to analyze, categorize, and label privacy risk statements in the policies of Android apps. The system uses a deeply tuned Distil- BERT transformer model that has the ability to comprehend legal terms and identify privacy-related semantics with high precision. The backend written in FastAPI works with text or URLs, and in real-time it makes inferences related to policy segments related to sharing data, location tracking, behavioral profiling, and sensitive information usage. A special policy-fetching module will automatically fetch and analyze recent app policies of such platforms as the Google Play Store. A browser extension can be used to protect users by protecting in real time as a privacy fil- ter, determining when an app is installed and displaying an intelligent Block, Warn, or Allow. It is a unified AI-driven system that enhances the user privacy awareness, compliance tracking, and how transformer-based NLP can enable users to use mobile securely and to maintain transparent data controls.

**Keywords:** Privacy policy analysis, Android security, transformer-based NLP, DistilBERT model, real-time risk detection, data privacy compliance, browser-extension monitoring

## 1 Introduction

The quick growth of Android app has transformed the digital interaction and made it easy to access services that may be in the form of banking or healthcare to communication and entertainment. But, as this expands, apps continuously gather personal information like location, contacts, history of use, and preference of use. Privacy policies, which are meant to enlighten the user about such practices, are in most cases long, legal and hard to comprehend. This makes it complicated to the extent that users cannot fully


B. Singh et al. (eds.), *Proceedings of the International Conference on Advances in Computing Technology and Artificial Intelligence (COMPUTATIA 2026)*, Atlantis Highlights in Intelligent Systems 18,
https://doi.org/10.2991/978-94-6239-713-2_16

understand the manner in which their data is gathered, stored and distributed. Consequently, an increasing demand exists to have automated systems that can help examine privacy policies and draw specific and actionable conclusions to protect the right of privacy of users. The lack of transparency in data-collection effects and unauthorized sharing of information is a growing threat on user trust in mobile ecosystems. According to the surveys, the majority of users do not read or misinterpret the privacy policies because of their legality and technicality. Cyber-attacks, malicious third-party trackers and data-misuse cases have increased the concerns, which highlights the issue of privacy-conscious decision making. The motivation behind this project is the necessity to empower users to have real time visibility of the privacy risks they expose to on an app level so that they can make informed choices regarding the installation and use of the app based on intelligent, automated policy interpretation. Despite the existence of several privacy analysis tools, the majority of them use manual analysis and/or static rule-based methods, as well as are confined to particular datasets and offline analysis. Current solutions are not real-time, cannot be very accurate in terms of legal text interpretation, and lack automated risk categorization and actionable installation-level advice. Additionally, available literature is mainly centered on English-based data and does not consider multilingual policy situations as found in international applications stores. This gap indicates the need of a transformer-based real-time privacy risk detection system that is incorporated within the application marketplaces.

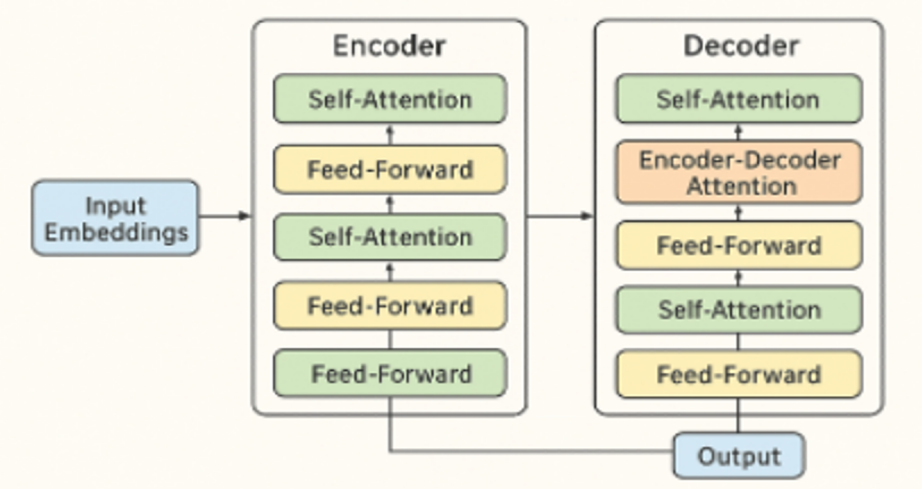


**Fig. 1.** Transformer based NLP

The language of privacy policies is very ambiguous, domain-specific and legal in nature and it is not easy to interpret it automatically. Determining sensitive clauses, distinguishing between compliant and risky statements and proper classification are not non-trivial tasks. Live risk evaluation also requires policy extraction and deep-learning deployment optimized on performance and necessitates a reliable extraction of policy out of dynamic sources, such as application stores. Also, it is challenging to provide the

risk outcomes to end users in a form understandable and practical, so balancing between technical accuracy and practical readability is necessary. This project aims at creating a natural language processing privacy threat detector framework in real-time on Android applications based on transformers. The system optimizes a DistilBERT model to help analyze the content of privacy policies, classify risks and/or classify apps as Block, Warn, or Allow. It has a FastAPI back-end, automated privacy policy extractor, and a browser extension that is embedded into the interface of the Google Play Store. The answer is restricted to a text-based analysis of privacy policies and English language content, though it lays the grounds of future implementation with multi-lingual support, contextual risk scoring and cross-platform implementation in the web and mobile ecosystem.

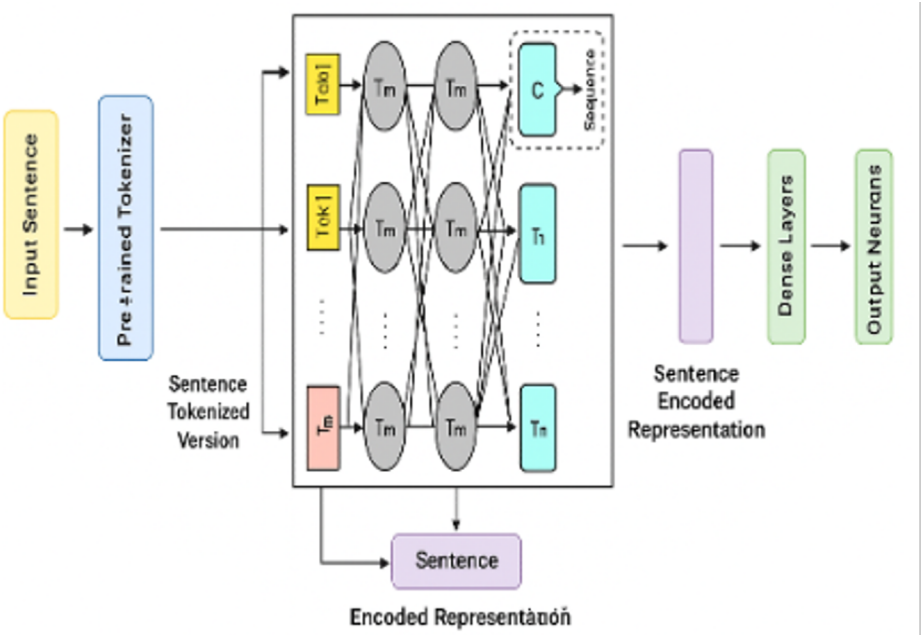


**Fig. 2.** DistilBERT Model Architecture

## 2 Literature Survey

A framework called Icon Checker that is meant to identify malicious activities in Android applications through the analysis of inconsistency between the icons and their actual behavior. The system detects rogue UI elements and hidden access of data with the help of icon-semantics classification and behavior analysis through machine learning. It has been demonstrated experimentally that it is very accurate in identifying malicious network communication hidden to users, thereby increasing user confidence and application openness [1]. A detailed system to examine the adversarial threats to bio-

metric systems such as fingerprints, facial and voice authentication. The research designs vulnerability parameters and deconstructs adversarial simulated samples to assess the system resilience. The results indicate that biometric systems are extremely vulnerable to spoofing, perturbation attacks, and deepfake-based manipulation. The study emphasizes defensive strategies against attacks including multimodal fusion and adversarial training of secure biometric recognition [2].

Privacy threats posed due to improperly set analytics services embedded in mobile apps. It demonstrates that unsecure API environments, overabundance of data gathering, and insufficient endpoint protection tend to reveal user data accidentally. Configuration flaws are determined in a large set of Android data via automated scanning methods. Findings indicate that misconfiguration is one of the significant privacy threats, and it often has an even higher impact on data leakage than deliberate malware [3]. The new threats in the LLM-based app store ecosystems, such as timely injection and unsafe execution of the plug-ins. Empirical testing demonstrates the weaknesses of the review mechanisms, sand- boxing and enforcement policies, which allow the malicious AI-driven attacks. Live threat simulation Live threat simulation evidences exfiltration of data, unsecured content generation, and user query manipulation. The article supports the idea of protecting AI marketplaces with secure execution layers, trust scoring and active monitoring [4].

A safe remote knowledge sharing system that safeguards confidential information by means of encryption, differentiated privacy, and federated control. It provides privacy measures in form of multi-layered privacy to maintain confidentiality of the communication patterns and user identities. Low latency and high privacy are confirmed through experimental evaluation which is good enough to support enterprise and research collaboration [5]. Malware detection based on machine learning is also the focus of adversarial attacks, which emphasize the techniques of feature manipulation and adversarial obfuscation. The updated malware can avoid classifiers through the simulation of harmless actions or injecting adversarial perturbations. The article analyzes such defense mechanisms as adversarial training, hybrid detection systems, and feature- robust models. It concludes that model-hardening and red-teaming initiatives in continuous fashion are necessary to enhance cybersecurity based on ML [6].

The obstacles that scientists encounter when trying to meet the standard of GDPR, but at the same time, working in large-volume data-driven research. The issues of strict consent, data-minimization, and deletion rights present an obstacle in the long-term use of datasets. Case studies provide insights into the balancing of privacy and research requirements using pseudonymization, controlled access, and reviewing the research ethics. The paper is demanding that more regulatory guidelines be followed to enhance innovation without infringing on the rights of personal privacy [7]. An audit log system that is based on blockchain and ensures immutability, transparency, and access control to comply with GDPR requirements. The architecture ensures the log records are secured, privacy is maintained and smart contracts are used to get consent. Its distributed structure makes it traceable without being tampered with, and the access to the sensitive data is limited to the authorized persons. Through experiments, there is an improvement in accountability and scalable audit performance, which fits regulated industries [8].

The perception and appreciation of privacy through the transparency of the consent messages presented in line with the principles of GDPR. Experiments on behavior reveal that users have more trust in a platform when secrecy options are pointed out in privacy notices. Results show that the readiness to share data are higher with the accent on transparency and control. The paper highlights the significance of privacy communication that is user-centric to improve trust and compliance [9]. The system traces the legal obligations to the software patterns to detect the insecure data handling and consent violations. The tool has been tested on large amounts of app data, and its recognition of privacy-related issues and frequent developer errors is quite high. It shows the significance of automated privacy checks during the initial stages of mobile development [10].

## 3 Problem Statement

Nevertheless, with the prevalence of Android applications, the general population does not know a lot about how their personal data is gathered, processed, and distributed because their privacy policy is quite time consuming, technical, and written in legal terminology. Consequently, users often accept granting permission without even being aware of the privacy implications, which creates a threat of sharing the data unauthorized, profiling behaviors, tracking by third parties, and the lack of adherence to privacy laws, such as GDPR and CCPA. The existing privacy analysis tools are mainly based on manual inspection and ad hoc rule-based tools which are not efficient and do not support real time decision-making and offer actionable advice during the time of app installation. Thus, an urgent requirement is the creation of an automated and clever system able to correctly analyze privacy policy, identify high-risk statements, and provide clear and real-time feedback on privacy awareness to users, allowing them to make informed decisions about their usage of apps and enhance the security of personal data within the Android system.

## 4 Existing System

Along with Android platforms, traditional privacy-risk detection methods mostly rely on the analysis of the static permissions, malware signature, or network surveillance to detect suspicious application practices. Nevertheless, these techniques fail frequently when malicious applications camouflage malicious activities in legitimate user actions and user interface transactions. In order to overcome this weakness, one of the outstanding current solutions is IconChecker which proposes a GUI-based anomaly detector system that can detect misleading icon behavior schemes in Android applications.

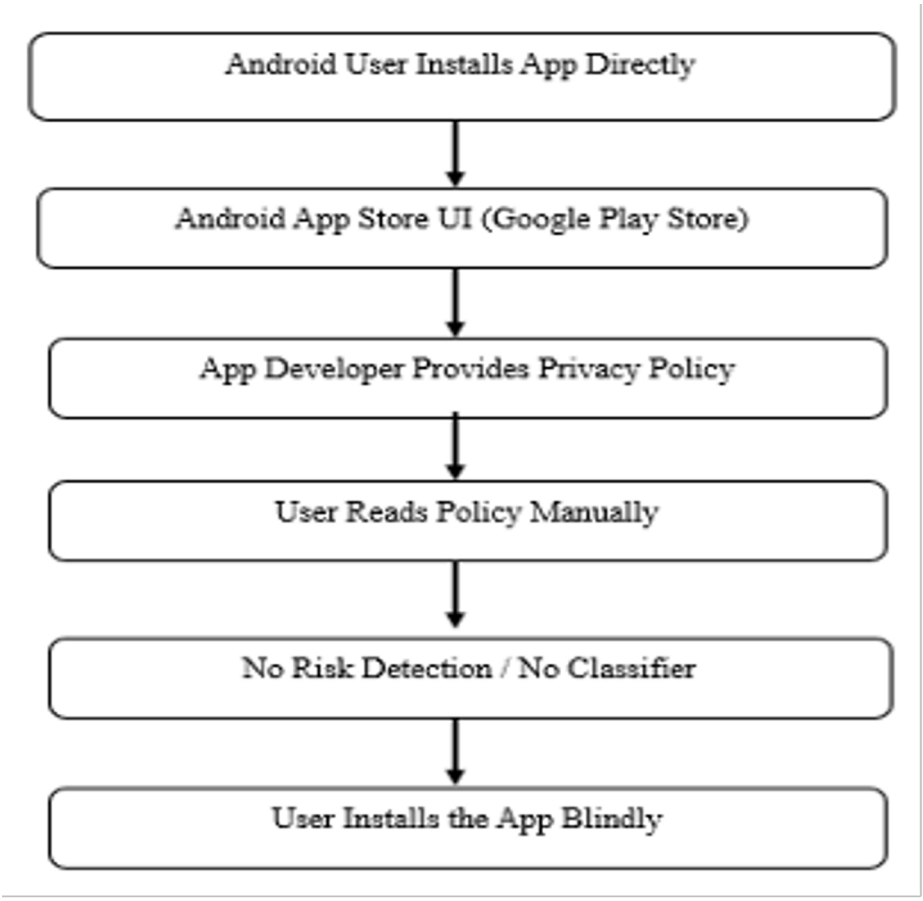


**Fig. 3.** Existing Block Diagram

Through tracking of user triggered actions and the reciprocating backend activity, it is very effective in identifying icons that are secretly executed to trigger malicious network actions like illegal transmission of data. This system is more precise in identifying the presence of malicious behavior associated with the use of a UI and is more effective than various baseline tools. Although this has been achieved, IconChecker primarily focuses on GUI abnormalities and subtextual patterns of communication; it does not examine the contents of any privacy pol- icy, lawful standards, or the clarity of data-handling declarations. Thus, despite the already available studies that enable lowering the threat analysis to behaviors, the privacy-policy interpretation that must be performed in real-time, as well as automated compliance evaluation-gaps that our transformer-based privacy-risk detector resolves are indeed critical.

# 5 Methodology

## 5.1 Collection and Preparation of Data Set

The starting point of the methodology is the gathering of various privacy policy documents by several sources that can be trusted such as Google Play Store applications, Apple App Store entries, and publicly available legal policy repositories. These policies are automatically collected and maintained into a rich set of policies intersecting many fields of application, like finance, healthcare, social networks, and utilities. After collection, the text is purged to eliminate ads, navigational elements and HTML elements. It then breaks down the policy text into substantial clauses and sentences since the analysis of privacy policy needs to know the context at the level of the clauses. All the clauses have been manually categorized as data sharing, tracking by a third party, sensitive data usage and statements that are friendly to compliance, and this is used as ground truth to train.

## 5.2 Preprocessing of Text and Processing of Features

The privacy policy text is preprocessed after being prepared as a dataset. These involve tokenization, elimination of stop words where needed and parsing of legal phrases into a machine-readable format. They are given special tokens to represent URLs, emails and numbers so that they are uniformly represented. NLP parsing rules are used to perform clauses segmentation without losing technical legal semantics. Other linguistic characteristics like the length of the sentences, frequency of key words and contextual indicators are also extracted so as to increase the awareness of the model to privacy-critical language patterns.

## 5.3 Transformer-based NLP Model Training

A delicate-tuned DistilBERT transformer model is used to classify privacy risks on a clause level. The processed text is inputted into the model, and the contextual significance along with legal jargon are learnt by undergoing supervised training. The labels are used to optimize the model, and the hyperparameters are adjusted to provide a high accuracy, precision, recall, and F1-score. DistilBERT model provides a classification label as well as a risk confidence score on each clause. In the process of training, continuous monitoring of evaluation metrics is implemented to avoid overfitting and to keep the performance across unseen privacy policies.

## 5.4 Backend Integration and Real-Time Inference

After the model has been trained, it is used in a FastAPI backend architecture to allow real-time privacy analysis. The input to the backend can take many different forms, which include raw text, PDF material, or policy links in the form of a URL. When provided exclusively with the application link, a policy fetching module can retrieve

live documents on privacy directly in the listings of the application in the app stores. The backend processes the text and runs the trained model and provides organized results such as clause and general privacy risk scores. The mechanisms of error handling, API security and latency optimization assure the rapid, reliable and secure inference.

### 5.5 Browser Extension and User-End Deployment

A small extension browser is created that functions as a live privacy protector and is mostly incorporated into the Google Play store interface. The extension sends the app URL to the backend when users see or even when they are trying to install an application. The privacy policy is evaluated with the backend, and a recommendation is returned in the form of Block, Warn, or Allow. The user is displayed with this choice on a clean visual interface through color coded badges and small explanations so he or she can make an informed decision without going through big documents.

### 5.6 System Evaluation and Validation

The system is also extensively tested on various measures such as accuracy, precision, recall, F1-score, and inference time. There is user-based testing, which is carried out to determine the level of clarity, usability and effectiveness in providing guidance on installation decisions. The strength of the model is considered in comparison with the complicated and vague privacy language in order to pro- vide high reliability. The retraining cycles and periodical updates are included in order to ensure that the model is in tandem with the changing policy standards and regulatory trends.

## 6 Proposed System

The suggested system presents a smart and automated Real-Time Privacy Risk Detector of Android Applications that uses a powerful transformer-based natural language processing to understand privacy policy documents and identify possible risks of data-handling. A model based on DistilBERT is fine- tuned to comprehend legal and privacy-related terms, single out the sensitive policy statements, and categorize the evaluated risks into practical groups. The system includes a FastAPI based real-time inference, enabling users to make real-time privacy assessments of text or URLs they have entered. Also, a policy- fetching engine is specifically designed to fetch live privacy policies of the Google Play Store and provide updated and correct analysis. A browser extension goes an extra step further to make it very easy to use, as it scans app policies during the installation phase and displays them as Block, Warn, or Allow, thus helping the user towards making safe app-selection decisions. This is an end-to-end auto- mated system that enhances privacy transparency, reduces data-security issues,

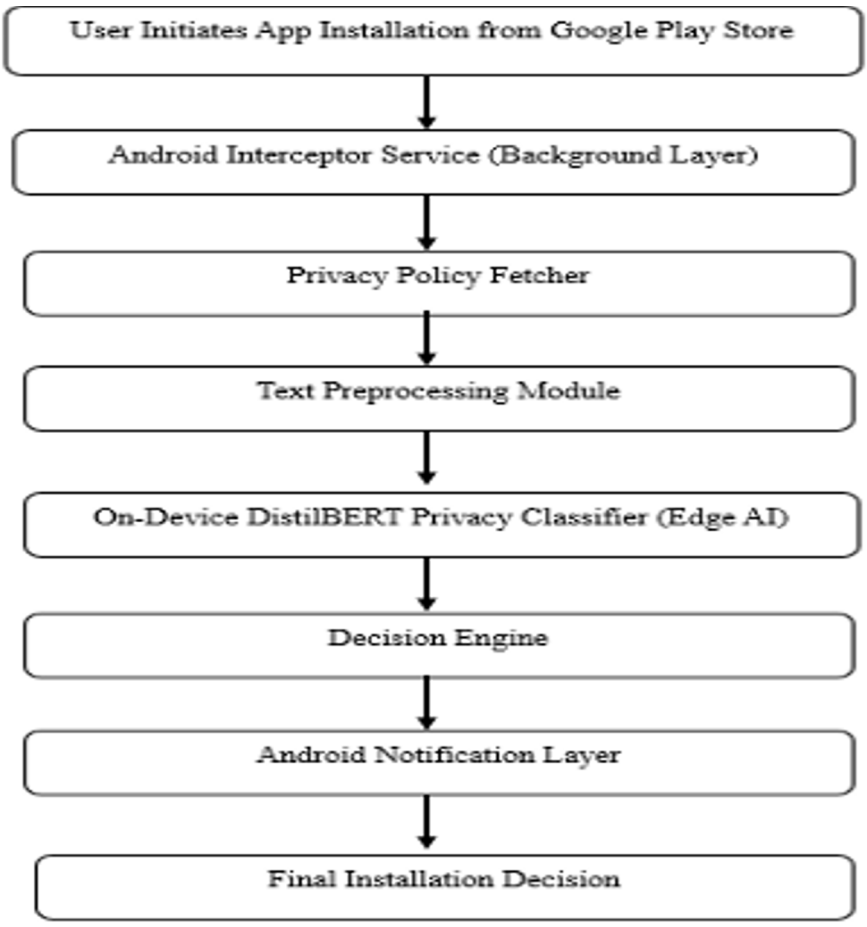


**Fig. 4.** Proposed Block Diagram

and provides the users with instant, relevant information on whether the apps are compliant with privacy. The suggested system will fill the gap between detailed privacy documents and user awareness to generate safer online interactions and encourage responsible data-governance behavior.

## 7 Result and Discussion

The Real-Time Privacy Risk Detector implementation starts by the creation of an annotated dataset of clauses of privacy policy extracted out of Android apps. These policies are just gathered on the pages of Google Play Store and divided by hands to meaningful statements. The clauses are then marked with the risk categories creating a benchmark training corpus. The data pipeline contains preprocessing steps, which are text normalization, clauses-level tokenization, and special encoding of tokens (URLs

and sensitive identifiers), and it makes sure that the input is all structured in the same way to the model. The DistilBERT model is then fine-tuned with the aid of the supervised

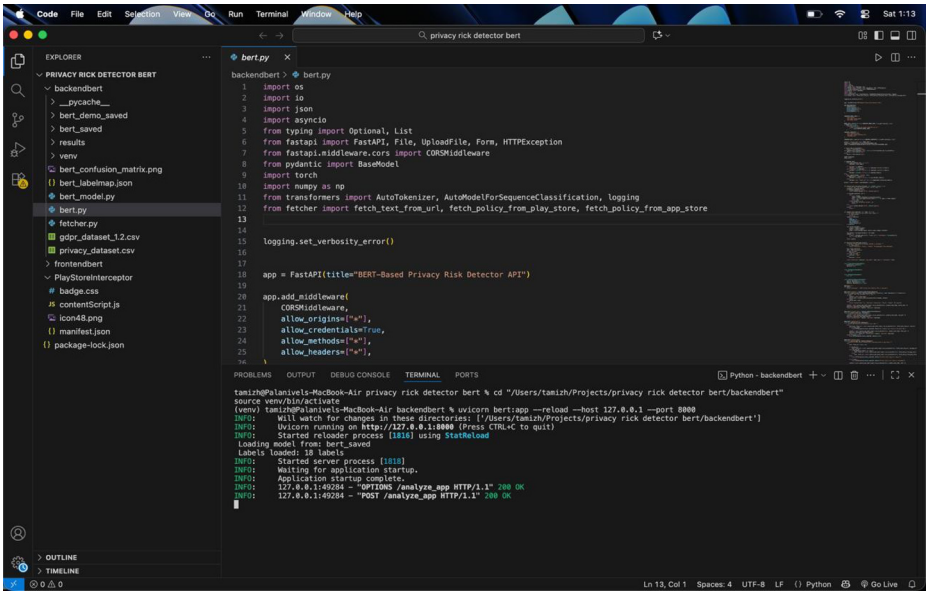


**Fig. 5.** Application Running

learning applied to the ready dataset. The implementation of the training process is done with the help of Python and the PyTorch-based transformer libraries and with the help of such methods as the learning-rate scheduling, the early stopping, and the class-balanced loss optimization. The purpose is to train the model so that it would comprehend legal terms and would recognize privacy- risk factors. In the development of a model, such performance measures as accuracy, F1-score, and validation loss are tracked to prevent overfitting and guarantee generalization. The system is trained and deployed to a FastAPI server in order to evaluate privacy in real-time. It has a dynamic policy-fetching module that will get live privacy policy text on app pages and inference is made immediately upon request. Moreover, a Chrome browser extension can be used to drive analysis when trying to install an app, and it will show clever suggestions as Block, Warn, or Allow. This deployment structure provides a smooth user interaction and the effective privacy risk identification without reading documents manually.

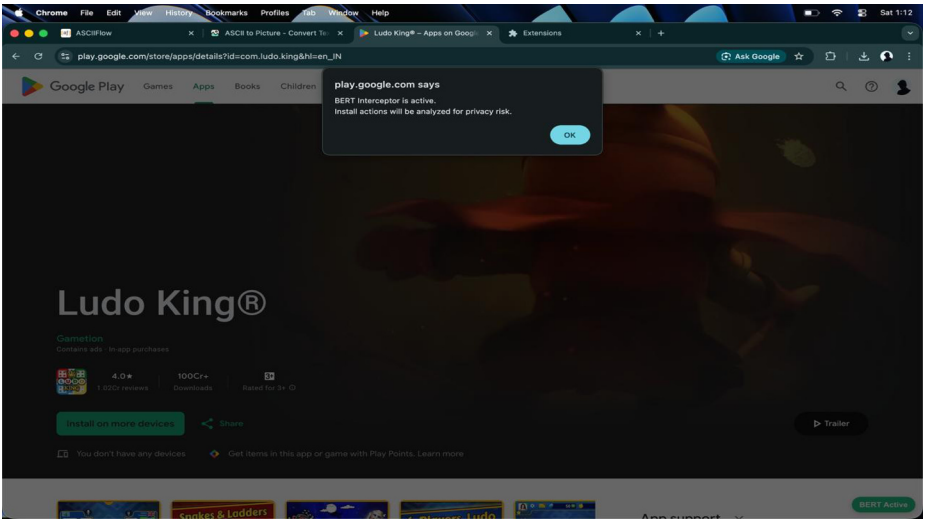


**Fig. 6.** Result

The obtained, trained DistilBERT-based privacy detector has high accuracy in detecting and classifying incorrect privacy-risk clauses. The model scores high evaluation results, and competitive precision and recall scores in the following category data sharing, third-party access, location tracking, and sensitive-data usage. The analysis of the confusion matrix verifies the strength of the model to differentiate the benign statements and high-risk content so that the majority of privacy-sensitive clauses can be correctly identified. During real-time inference tests, the system offers privacy evaluation in milliseconds, which is satisfactory with the objective of low-latency policy analysis. The browser extension effectively captures clicks to install apps and reads policy information in real- time Play Store listings and presents timely notifications to users. Throughout the test with sample applications in the fields of finance, healthcare, gaming, and utility, the system was able to highlight the concealed data-sharing pattern, fuzzy legal provisions, and showcase privacy-infringing patterns that are usually neglected by end-users. According to the user input tests on the prototype, the risk badges and summary explanations are really helpful in making a decision. The users complained that they felt more aware of the risks of privacy and that they were more ready to stop installing risky applications. The simplicity of color-coded

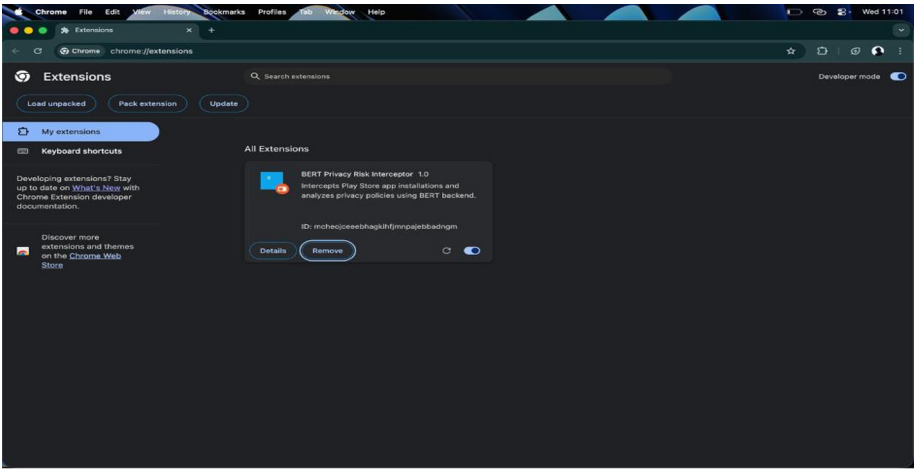


**Fig. 7.** Extension

choices and emphasized risk phrases have added to the high-usability scores, which have proved the usefulness of the system in the real life. The findings confirm that NLP using transformers is very useful in the interpretation of privacy policies and automated risk evaluation. DistilBERT model has shown great knowledge in the legal language, which is better than the traditional keyword- based and rule-based methods. The capacity to identify subtle utterances and connoted information-sharing purpose makes the system an effective contender of real-time privacy surveillance. The policy-fetching mechanism helps in ensuring that an assessment is done using up-to-date content of the app-store that would solve a significant weakness of the current systems that depend on fixed datasets.

The system performance, however, depends on the complexity of the policy language; some vague or deliberately ambiguous terms question the accuracy of interpretation of the model. There is no support to use multilingual policies and policy documents in PDF or image formats need to be manually extracted. The challenges indicate the future work directions of multilingual support, OCR- based extraction, transformer fine-tuning in cross-domain legal understanding and the combination of permission-analysis tools to gain extra information on privacy.

## 8 Conclusion

The framework proposed is able to adequately respond to increased demand of transparent, automated and real-time privacy-risk evaluation in Android applications. The system will give users a clear understanding about the way their personal information is being gathered, processed and distributed by mobile applications by combining a DistilBERT-based transformer model, with real-time scraping and analyzing of browser-extensions. The solution is a viable solution to the discrepancy between complicated legal privacy policies and the capability of end-users to comprehend and assess

them to achieve well-informed consent and increased digital autonomy. The methodology has proven to be robust in terms of experimental results showing good performance in terms of classification accuracy, speed of detection and practical usability. The contribution of the work is a scalable and intelligent privacy compliance assistant that can assist in safer use of the apps, enhance user trust and make digital service providers more accountable. Additions to this in the future can be made with support of multiple languages, integration with other platforms and better threat-detection models to enhance user privacy in developing digital ecosystems.

## References

1. Li, Y., Feng, R., Chen, S., Guo, Q., Fan, L., Li, X.: IconChecker: Anomaly detection of icon-behaviors for Android apps. In: Proceedings of the 28th Asia-Pacific Software Engineering Conference (APSEC 2021). IEEE (2021)
2. Park, S.H., Lee, S.-H., Lim, M.Y., Hong, P.M., Lee, Y.K.: A comprehensive risk analysis method for adversarial attacks on biometric authentication systems. IEEE Access (2024)
3. Zhang, X., Wang, X., Slavin, R., Breaux, T., Niu, J.: How does misconfiguration of analytic services compromise mobile privacy. In: Proceedings of the IEEE/ACM 42nd International Conference on Software Engineering (ICSE 2020). IEEE/ACM (2020)
4. Hou, X., Zhao, Y., Wang, H.: On the (in)security of LLM app stores. In: Proceedings of the IEEE Symposium on Security and Privacy (SP 2025). IEEE (2025)
5. Dahlmanns, M., Dax, C., Matzutt, R., Pennekamp, J., Hiller, J., Wehrle, K.: Privacy-preserving remote knowledge system. In: Proceedings of the IEEE 27th International Conference on Network Protocols (ICNP 2019). IEEE (2019)
6. Qi, X., Tang, Y., Wang, Y., Liu, T., Jing, J.: Adversarial example attacks against intelligent malware detection: A survey. In: Proceedings of the 4th International Conference on Applied Machine Learning (ICAML 2022). IEEE (2022)
7. Meszaros, J.: The conflict between privacy and scientific research in the GDPR. In: Proceedings of the Pacific Neighborhood Consortium Annual Conference and Joint Meetings (PNC 2018) (2018)
8. Aslan, U., Şen, B.: GDPR compliant audit log management system with blockchain. In: Proceedings of the 15th Turkish National Software Engineering Symposium (UYMS 2021) (2021)
9. Annam, J.R., Ande, P.K., Kanuri, B., Prasad, C., Babu, B.S., Tatineni, P.: User valuation of secrecy framing based on GDPR users. In: Proceedings of the Third International Conference on Inventive Research in Computing Applications (ICIRCA 2021). IEEE (2021)
10. Khedkar, M.: Static analysis for Android GDPR compliance assurance. In: Proceedings of the IEEE/ACM 45th International Conference on Software Engineering: Companion Proceedings (ICSE-Companion 2023). IEEE/ACM (2023)
11. Niya, S.R., Willems, J., Stiller, B.: A case study of a blockchain-GDPR adaptation. In: Proceedings of the IEEE International Conference on Blockchain and Cryptocurrency (ICBC 2022). IEEE (2022)
12. Magoulas, G.S., Polykalas, S.E.: Access to personal data is still tempting for mobile apps even after the GDPR implementation. In: Proceedings of the 8th South-East Europe Design Automation, Computer Engineering, Computer Networks and Social Media Conference (SEEDA-CECNSM 2023). IEEE (2023)

13. Sağlam, R.B., Aslan, Ç.B., Li, S., Dickson, L., Pogrebna, G.: A data-driven analysis of blockchain systems' public online communications on GDPR. In: Proceedings of the IEEE International Conference on Decentralized Applications and Infrastructures (DAPPS 2020). IEEE (2020)
14. Kulkarni, M.S., Naik, H.L., Bharathi, S.V.: Textual analysis of privacy policies to understand the effect of GDPR. In: Proceedings of the 2nd International Conference on Futuristic Technologies (INCOFT 2023) (2023)
15. Dave, A., Agrawal, A.K.: A comparative study with GDPR, HIPAA, CCPA, PIPEDA and DPDPA. In: Proceedings of the IEEE International Conference on Computer, Electronics, Electrical Engineering and their Applications (IC2E3 2025). IEEE (2025)
16. Human, S., Pandit, H.J., Morel, V., Santos, C., Degeling, M., Rossi, A., Botes, W., Jesus, V., Kamara, I.: Data protection and consenting communication mechanisms: Current open proposals and challenges. In: Proceedings of the IEEE European Symposium on Security and Privacy Workshops (EuroSPW 2022). IEEE (2022)
17. Saini, H., Prasad, S.K.: Machine learning approaches to securing IoT networks: A comprehensive review of recent advances. In: Proceedings of the 2nd International Conference on Advancements and Key Challenges in Green Energy and Computing (AKGEC 2024) (2024)
18. Varghese, S.R., Juliet, S., N.S., A.: Social media text analysis for disaster management using DistilBERT model. In: Proceedings of the International Conference on Science Technology Engineering and Management (ICSTEM 2024) (2024)
19. Prema, V., Elavazhahan, V.: Sculpting DistilBERT: Enhancing efficiency in resource-constrained scenarios. In: Proceedings of the 12th International Conference on System Modeling Advancement in Research Trends (SMART 2023) (2023)
20. Gakpetor, J.M., Doe, M., Damoah, M.Y.S., Damoah, D.D., Arthur, J.K., Asare, M.T.: AI-generated and human-written text detection using DistilBERT. In: Proceedings of IEEE SmartBlock4Africa 2024. IEEE (2024)