Pinocchio

Scam Detection App

https://youtu.be/4sU-ZUZL2wo
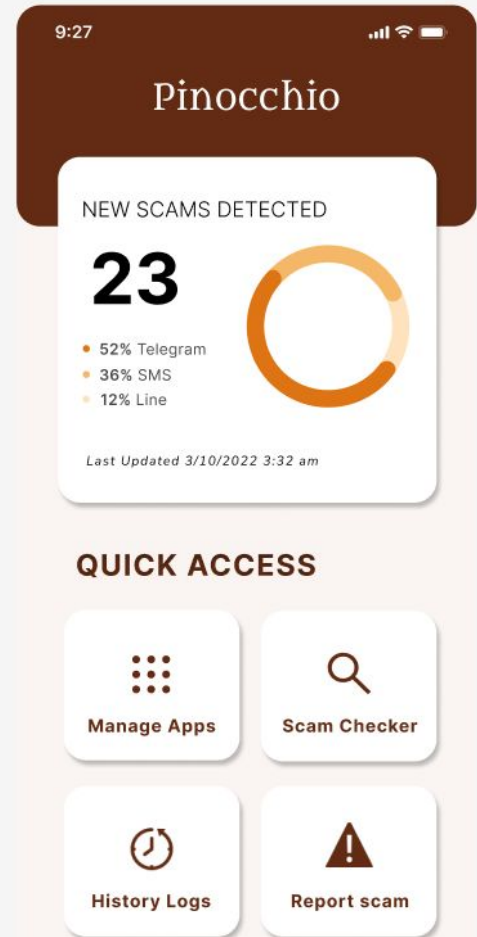
The Problem

As Singapore transforms into <mark>smart digital nation</mark> through technology, it becomes important to shield people from the **increasing number of text scams.**

Scammers are also **becoming smarter** and **more deceptive**

THE STRAITS TIMES — SINGAPORE — LOG IN

Nearly $1 billion lost by scam victims in Singapore since 2016

cna | Top Stories | Latest News | Discover | Singapore | Asia | Commentary | Sustainability | CNA Insider | Lifestyle | Watch | Listen | + All Sections

Singapore

'I just felt really stupid': How a scam victim almost lost more than S$1 million – until police and the bank intervened

THE STRAITS TIMES — SINGAPORE — LOG IN

At least 587 victims lost $2.7 million to scammers posing as friends since start of 2022

cna | Top Stories | Latest News | Discover | Singapore | Asia | Commentary | Sustainability | CNA Insider | Lifestyle | Watch | Listen | + All Sections

Singapore

325 suspects investigated over scams involving more than S$9 million

The 201 men and 124 women are aged between 16 and 86.

THE STRAITS TIMES — SINGAPORE

Scam victims in S'pore lost $633.3 million in 2021

The Solution

An **app** which monitors messaging platforms and consolidates scam or spam chats for the user to **sort, look through or delete**.

Market Research

## Similar Apps
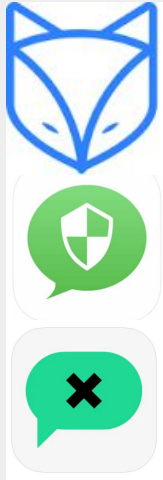
ScamShield
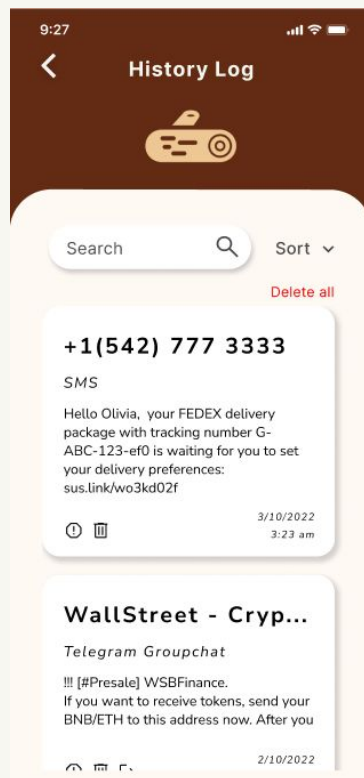
SMS Shield

TextKiller

Our Unique Selling Points

- Scam detection across multiple messaging platforms (other than SMS)
- Compatible on both Android and iOS devices
- More user options (Delete / Leave / Sort)
- Manual scam message checker

The Features

Consolidates scam chats across different messaging platforms

Delete / Leave / Sort compiled chats
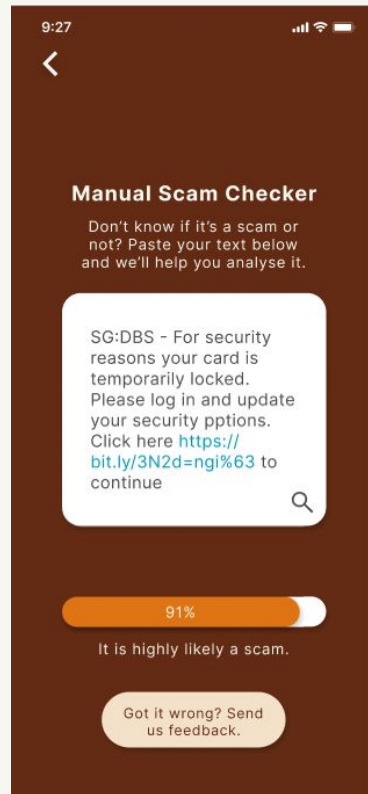
Textbox for users to manually check if a message is a scam

Natural Language Processing for scam chats detection

Incorporating Artificial Intelligence

- **Natural Language Processing (NLP)**
  powers the spam message detection feature by predicting the likelihood of the
  message received being a scam

| Training the ML Model with existing scam dataset | Data undergoes preprocessing (i.e. Tokenization, Stemming) | Wrangled data trains Random Forest Classification Model | New input data is entered into ML Model for prediction | AI Model determines whether or not the message is a scam |

Potential Issues & Solutions

Privacy Issues
Pinocchio does not store any user data unless permitted to

Inaccurate Detection
Users may report any inaccurate detections to help train the AI further

No access to most platform's delete or leave chat function (not all platforms have ready-to-use libraries like Telegram & Telethon)
Partner up with them to gain more access to those functions

Scalability & Future Considerations

- **Increasing number of supported platforms**
  Requires AI Model to be updated regularly due to the increase in different types of scams

- **Increasing accuracy of scam detection**
  Adding deep facial recognition for profile pictures of scammers (usually similar), voice to text analysis to detect scam phone calls, web scraping of website data to detect scam sites, and AI reinforcement learning

- **Collecting scam data to help in anti-scam efforts**
  Collecting certain data such as the profile picture, number, message(s) sent, etc, of the chat may help the government fight scams (SPF)

- **Monetisation**
  Offer a free Basic Plan (SMS & WhatsApp only) and a subscription-based Premium Plan (All compatible platforms, including calls)