1. Giới thiệu
   1. Giới thiệu đề tài
   2. Mục tiêu và phạm vi đề bài
2. Phân tích vấn đề và hướng giải quyết
   1. Nhận diện giọng nói
      1. Vấn đề
      2. Giải pháp
   2. Độ chính xác của các api
      1. Vấn đề
      2. Giải pháp
   3. Âm điệu, âm vực khác nhau.
      1. Vấn đề
      2. Giải pháp
   4. Xử lí tiếng ồn của dữ liệu đầu vào
      1. Vấn đề
      2. Giải pháp
   5. Quản lý dữ liệu, người dùng
      1. Vấn đề
      2. Giải pháp
   6. Ngôn ngữ hỗ trợ
   7. Lưu trữ báo cáo
   8. Nhận diện dấu chấm, dấu phẩy
   9. Nhận diện chữ số
   10. Các thông số quan tâm
       1. Vấn đề

Do ứng dụng dược xây dựng trên nền tảng Smartphone nên có những vấn đề về tốc độ, hiệu năng, bộ nhớ cần phải được quan tâm. Đa số các thiết bị chạy hệ điều hành Android đều là các thiết bị có tốc độ, bộ nhớ nhỏ hơn nhiều lần so với các máy tính. Vì vậy việc tối ưu hóa tốc độ, hiệu năng và bộ nhớ khi sử dụng ứng dụng là một vấn đề quan trọng cần phải được quan tâm. Có rất nhiều những vấn đề cần phải quan tâm khi xây dựng ứng dụng Android như tốc độ xử lý, độ mượt, bộ nhớ, tiêu tốn pin,….

* + 1. Giải pháp

Do nhóm vẫn chưa xây dựng ứng dụng nên vẫn chưa đưa ra được những giải pháp cụ thể cho các vấn đề về hiệu suất của ứng dụng sẽ xây dựng. Một số vấn đề mà nhóm định sẽ quan tâm và cố gắng tối ưu là:

* Tốc độ xử lý của ứng dụng: Vì là một ứng dụng nhằm mục đích để tạo báo cáo nhanh nên tốc độ xử lý của ứng dụng là một vấn đề quan trọng. Nếu mất quá nhiều thời gian để xử lý thì sẽ ảnh hưởng để công việc, mức độ xử dụng của người dùng.
* Mức độ tiêu tốn pin: Ứng dụng xây dựng trên Smartphone nên nếu tiêu tốn quá nhiều pin khi sử dụng thì sẽ dẫn đến việc thời gian sử dụng ứng dụng ngắn. Người dùng khó có thể hoàn thành công việc của mình.
* Độ mượt của ứng dụng khi sử dụng.
* Dung lượng, tiêu tốn bộ nhớ khi sử dụng: Do smartphone thường có bộ nhớ nhỏ, RAM thấp nên cần phải tối ưu hóa về sử dụng bộ nhớ trong ứng dụng đề tránh tình trạng bị tràn bộ nhớ khiến thiết bị bị frozen.
  1. Bảo mật thông tin
     1. Vấn đề

Ứng dụng định xây dựng là một ứng dụng hướng đến người dùng là các doanh nghiệp. Các báo cáo, các dữ liệu mà người dùng nhập vào đều là những thông tin bí mật của công ty. Do đó các thông tin, dữ liệu được lưu trữ và truyền đi đều phải được mã hóa để đảm bảo không bị những người thứ 3 đánh cắp thông tin.

* + 1. Giải pháp

Nhóm dự định sẽ sử dụng giải thuật mã hóa Advanced Encryption Standard (AES) để bảo mật dữ liệu. Sau khi ứng dụng xủ lý và chuẩn bị gửi dữ liệu lên server, dữ liệu đó sẽ được mã hóa bằng AES rồi sẽ được gửi lên lưu trữ trên server.

* + 1. Giải thuật Advanced Encryption Standard (AES)

AES là một trong những giải thuật mã hóa được sử dụng rộng rãi nhất hiện nay. Nó được phát triển bởi tổ chức National Institute of Standards and Technology của chỉnh phủ Mỹ, nhằm thay thế cho giải thuật Data Encryption Standard (DES). AES là một giải thuật mã hóa sử dụng khóa đối xứng. Có nghĩa là nó sử dụng một secret key cho cả việc mã hóa và giải mã.

* Độ an toàn của AES: AES được công nhận là rất đáng tin cậy khi sử dụng trong thực tế. Nó được sử dụng rộng rãi để mã hóa trong nhiều lĩnh vực, bao gồm cả chính phủ, các ứng dụng end-user,… AES cung cấp 3 độ dài khóa: 128-bit, 192-bit, 256-bit. Khóa có độ dài trên 192-bit đủ để bảo mật cả những văn bản tuyệt mật của chính phủ.
  1. Kết nối mạng
     1. Vấn đề

Ứng dung sử dụng các API để xử lý giọng nói, sử dụng Database Web Server để lưu trữ dữ liệu,... nên ứng dụng cần phải có kết nối Internet để có thể hoạt động được.

* + 1. Giải pháp

Nhóm quyết định điện thoại bắt buộc phải có kết nối đến Internet để có thể sử dụng được ứng dụng.

1. Các ứng dụng, công nghệ hiện tại
   1. Google Cloud Speech API

Google Cloud Speech API là một api được xây dựng bởi Google.

* Hổ trợ bởi công nghệ Machine Learning

API được áp dụng những thuật toán mạng neutral học sâu tiên tiến nhất để chuyển từ giọng nói của người dùng sang văn bản với độ chính xác cao.

* Hỗ trợ hơn 80 ngôn ngữ

API có thể nhận diện được hơn 80 ngôn ngữ và cách đọc khác nhau để hỗ trợ cho nền tảng người dùng toàn cầu.

* Trả về kết quả trong Real-Time

API có thể trả về kết quả một cách liên tục, ngay khi kết quả được nhận diện xong. Kết quả có thể xuất hiện ngay khi đang nói. API còn có thể trả về kết quả nhận diện từ giọng nói được lưu trữ dạng file.

* Độ chính xác trong môi trường nhiều tiếng ồn

Không cần tiền xử lí tín hiệu hoặc loại bỏ tiếng ồn trước gửi đến API. Dịch vụ của API có thể xử lý được tiếng ồn đến từ nhiều loại môi trường khác nhau.

* Từ gợi ý

Nhận diện âm thanh có thể được tùy biến thành một ngữ cảnh riêng biệt bằng cách cung cấp một tập các từ, câu mà có vẻ giống câu được nói.

* 1. IBM Watson Speech To Text API

Watson Speech To Text API cho phép người dùng có thể chuyển giọng nói thành văn bản. Để nhận diện và chuyển giọng nói một cách chính xác, API áp dụng trí tuệ nhân tạo để kết hợp thông tin về ngữ pháp và cấu trúc ngôn ngữ với thông tin từ tín hiệu âm thanh. API có thể liên tục trả về và cập nhật kết quả khi nhận được nhiều thông tin hơn.

* 1. Bing Speech API

Microsoft’s Speech Service chuyển giọng nói thành văn bản phù hợp để hiển thị cho người dùng. Văn bản được dịch bao gồm thểm viết hoa, dấu chấm câu, chuẩn hóa văn bản.... Microsoft’s Speech Services cung cấp 2 lựa chọn cho việc xây dựng app:

* REST API sử dụng chunked-transfer decoding để chuyển một câu ngắn mà không cần real-time streaming
* WebSocket API sử dụng full-duplex communication để chuyển một đoạn âm thanh dài và hỗ trợ hiện thị ngay lập tức kết quả.
  1. So sánh các API

1. Phân tích hệ thống
   1. Chức năng hệ thống

Các chức năng của hệ thống mà nhóm định xây dựng.

* + 1. Ứng dụng di dộng tạo báo cáo nhanh bằng giọng nói
* Tạo báo cáo theo mẫu chọn trước: Nhập dữ liệu vào bằng giọng nói, ứng dụng sẽ xử lý và hiển thị văn bản ra màn hình.
* Chỉnh sửa dữ liệu trong quá trình tạo báo cáo
* Chỉnh sửa báo cáo đã tạo
* Cho phép tạo mẫu báo cáo mới
* Xem những báo cáo đã tạo
* Tải về các báo đã tạo
* Quản lí tài khoản, đăng nhập, đăng kí.
  + 1. Web quản lý báo cáo và người dùng.
* Xem danh sách người dùng, danh sách báo cáo
* Chỉnh sửa báo cáo, thông tin người dùng
* Thêm, xóa người dùng
* Xóa báo cáo, in và tải về báo cáo
* Thống kê dữ liệu
* Đăng nhập
  1. Use Case Diagram
     1. Lược đồ Use Case
     2. Đặc tả Use Case
  2. Class Diagram

1. Tổng kết
   1. Các kết quả đã đạt được
   2. Hướng phát triển
   3. Tài liệu tham khảo

Cloud Speech API: cloud.google.com/speech/

Bing Speech API: www.azure.microsoft.com

IBM Watson Developer Cloud: www.ibm.com/watson