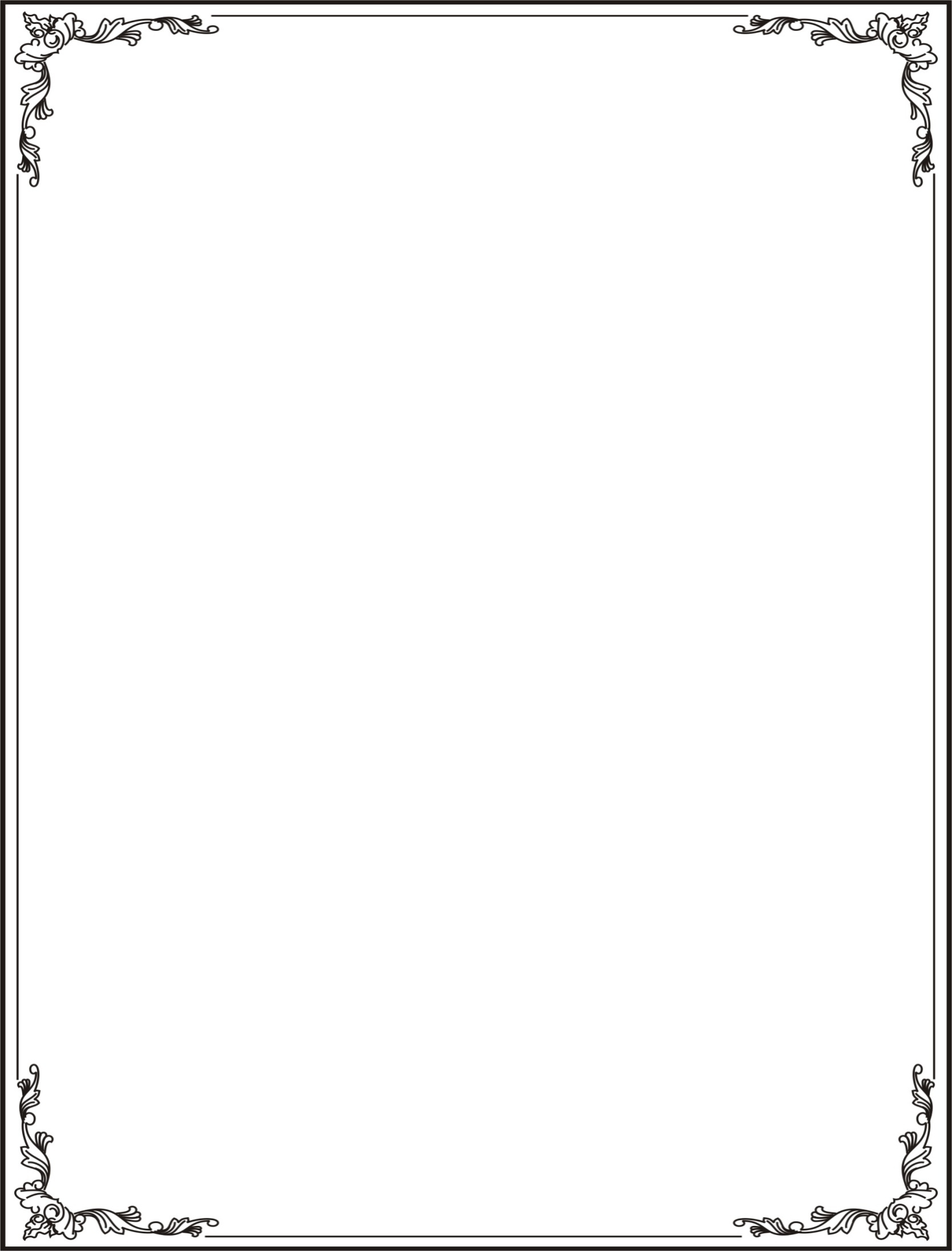
**ĐẠI HỌC QUỐC GIA THÀNH PHỐ HỒ CHÍ MINH**

**TRƯỜNG ĐẠI HỌC BÁCH KHOA THÀNH PHỐ HỒ CHÍ MINH**

**BÁO CÁO ĐỀ CƯƠNG LUẬN VĂN TỐT NGHIỆP**

**Phần : Nhận dạng giọng nói**

**Giảng viên hướng dẫn : Đặng Trần Trí**

**Phan Trần Thái Sơn - 1512853**

**Huỳnh Minh Thịnh - 1513245**

**Thống kê nhận diện giọng nói**

Sau đây là bảng thống kê của app nhận diện giọng nói khi sử dụng API Cloud Google Speech-to-Text

Trong đó :

* Câu nói : Câu nói do mình nói
* Đoạn Text : Đoạn text xuất ra từ app khi mình nói
* Chính xác : Độ chính xác của đoạn text và câu nói đó

|  |  |  |
| --- | --- | --- |
| **Câu nói** | **Đoạn Text** | **Chính xác** |
| 123 | 123 | 100% |
| Xin chào các bạn | Xin chào các bạn | 100% |
| A B C D E F G H M N K J | A B C D E F G H M N K G | 91.6% |
| Chào buổi sáng | Chào buổi sáng | 100% |
| Xin chào các bạn mình là Huỳnh Minh Thịnh | Xin chào các bạn mình là quyền Bính Thìn | ~80% |
| Cuộc sống mà nó khổ vậy đó | Cuộc sống mà nó khổ vậy đó | 100% |
| Phan Trần Thái Sơn là siêu nhân | Tranh trạng thái Sơn và siêu nhân | ~58% |
| Yasuo gank tem 15 gg | Yasuo gank team 15gg | >100% ( fix luôn cả tem thành team ) |
| Một con vịt xòe ra hai cái cánh | Một con vịt xòe ra hai cái cánh | 100% |
| SixMobile . Chất lượng làm nên thương hiệu | Xích Mobile chất lượng làm nên thương hiệu | 85.7% |
| Kim Dung là một trong những nhà văn ảnh hưởng nhất của văn học Trung Quốc hiện đại. Ông còn là người đồng sáng lập của nhật báo Hồng Kông Minh Báo, ra đời năm 1959 | Kim Dung là một trong những nhà khoa ảnh hưởng nhất của văn học Trung Quốc hiện đại . Không còn lòng người đọc Tiến Lộc của Nhật Báo Hồng Kông Minh báo cần năm 1900 năm | ~50% |
| 1 2 3 quẫy 1 2 3 quẫy 12 3 | 1 2 3 7 1 2 3 7 1 2 3 | 80% |
| Quyết tâm của ngài đúng là xứng đáng với vị thế một nhà lãnh đạo, nhưng mỗi **quyết** tâm thì không thể thay đổi được thế giới đâu | Quyết tâm của ngài đúng là xứng đáng với vị thế một nhà lãnh đạo, nhưng mỗi **mỹ** tâm thì không thể thay đổi được thế giới đâu | >95% |
| Điều này dễ thôi anh bạn , hãy đến với 6mobile | Chủ đề dễ thương bạn | 0% |
| Alo 1234 Alo 1234 Quẫy Quẫy nào | Alo 1234 alo 1234 vẫy vẫy vãy vãy vẫy | <50% |
| Chào buổi sáng bạn khỏe không ? | Chào buổi sáng bạn khỏe không ? | 100% |
| Đây là chiếc áo sơ mi và bên kia là cái áo gì đó | Đây là chiếc áo sơ mi và bên kia là cái áo gì đó | 100% |
| Trường Đại học bách khoa là trường giỏi nhất tại việt nam | Trường Đại học bách khoa **Hà Nội kim lân** là trường giỏ nhất tại việt nam | 90% |
| Tôi yêu thích anime và tôi yêu thích nhạc âu mỹ | Tôi yêu thích anime và tôi yêu thích nhạc âu mỹ | 100% |
| Ngày mai nghỉ học nha Sơn , tao lười học quá | Ngày mai nghỉ học nha **Sao** , tao lười học quá | >90% |
| (đọc nhanh) Tao buồn ngủ quá , tao đi ngủ đây | Tao buồn ngủ quá Tao muốn ngủ đi | >90% |
| (đọc nhanh )Coi bộ cái cửa cực kỳ chắc chắn, nhưng may một cái là lũ bên trong chạy lên tầng trệt đánh chặn tôi và quên khóa lại | Coi bộ **váy cưới** cực kỳ **chân trắng**, **ren may một cái là lũ bên trong chợ lên từ hạt phấn Trần Hương và hoa** | <30% |
| Tao biết nói gì bây giờ nhưng mà phải nói mà thôi | Tao biết nói gì bây giờ nhưng mà phải nói mà thôi |  |
| Liên Minh Huyền thoại là game dành cho đồng đội và chiến thật chứ không phải cá nhân **làm** nên tất cả | Liên Minh Huyền thoại là game dành cho đồng đội và chiến thật chứ không phải cá nhân nên tất cả | >98% |

**Phân tích**

**Qua 24 test case ví dụ , ta thấy :**

* Những đoạn text ngắn ( đọc chậm hay nhanh ) độ chính xác cực kỳ cao
* Những đoạn text cực kỳ dài độ chính xác đa số trên 50% ( khả quan ) tuy nhiên nếu đọc nhanh thì độ chính xác sẽ rất thấp
* Có thể nhận diện một chút tiếng anh mặc dù định nghĩa Tiếng Việt
* Có khả năng viết hoa tên người nào đó , thủ đô , hay ngành nào đó.

**Mục tiêu của Project** : Nhận diện giọng nói rồi sau đó xuất ra đoạn text để giúp người khiếm thính có thể hiểu mình đang nói gì

**Điều kiện Project** :

* Người nói phải nói không quá 1 phút cho 1 lần nhận diện giọng nói
* Giọng nói phải to và rõ , không cần nhanh
* Mạng 3G/Wifi khi sử dụng app phải tốt

**Chi phí sử dụng API Google Cloud :**

* Rẻ , được tặng 300$ cho tài khoản gắn visa để sử dụng dịch vụ google
* Khả năng chính xác cao , hệ thống không bị delay nhiều , có thể gọi nhận diện giọng nói nhiều lần mà không cần phải đợi 1 phút
* 1 đoạn text dài nhưng xử lý cũng khá nhanh và cũng cho kết quả khả quan

**KẾT LUẬN**

Sau khi phân tích đặc điểm , điều kiện , mục tiêu và chi phí sử dụng của API google cloud , nhóm chúng tôi quyết định sẽ dùng API Google Cloud để làm đề cương luận văn tốt nghiệp .

Để cho đề cương luận văn tốt nghiệp và luận văn tốt nghiệp tru trơn hơn , API Google Cloud là ứng cử viên sáng giá cho đề tài này và hi vọng với nó , sẽ giúp cho mọi người khiếm thính có khả năng giao tiếp với mọi người trong xã hội

Link GitHub Project :

https://github.com/Kurosaki12a1/SpeechToText