**BẢN BÁO CÁO**

**Nhóm 14**

**Đề tài: Đề xuất hệ thống học từ vựng thi Tiếng Anh.**

Nhóm em đã làm được những việc:

* Tách từ, loại bỏ stop words từ một văn bản sang 1 list.
* Sử dụng Bag Of Word để tính các giá trị tần số, vấn đề em gặp ở đây là sau khi tính thì không biết vận dụng nó thế nào và chưa hiểu ứng dụng của nó như thế nào dẫn đến còn phân vân về hướng giải quyết bài toán.
* Hướng giải quyết hiện tại của em là sẽ import các văn bản vào, sau đó tách từ và loại bỏ stop words, từ danh sách các từ đã có, tiếp tục tính số lần xuất hiện của mỗi từ và sắp xếp theo một danh sách giảm dần. Sau đó vẽ biểu đồ thể hiện mức độ xuất hiện của mỗi từ để đánh giá và xem xét.
* Em có suy nghĩ một số ý tưởng liên quan đến khảo sát và sử dụng kiến thức nghiên cứu môn Anh Văn nhưng em thấy không khả thi vì đề tài mình thiên về Machine Learning hơn.
* Kết luận lại vấn đề của em hiện tại là không biết hướng đi để giải quyết bài toán như thế nào, nên không biết phải làm gì hay học gì tiếp theo ạ. Vì thế nên tiến độ công việc của nhóm còn chậm so với dự định. Sau khi biết được hướng đi và được cô hướng dẫn, nhóm em sẽ hoàn thành công việc hiệu quả hơn ạ.
* Những kiến thức em học được và những gì em làm được thể hiện trong 2 file code sau ạ.