**BÁO CÁO TIẾN ĐỘ THỰC HIỆN ĐỒ ÁN TỐT NGHIỆP**

**ĐỀ TÀI: XÂY DỰNG WEBSITE DIỄN ĐÀN HỎI ĐÁP TRỰC TUYẾN VỀ CÔNG NGHỆ**

------------------------

**GVHD:** TS. Huỳnh Xuân Phụng

**SVTH:** Trấn Tấn Quý

Phạm Hoàng Việt

1. **PHẠM VI NGHIÊN CỨU**
2. **MỤC TIÊU**

Xây dựng website hỏi đáp trực tuyến về công nghệ với những tính năng:

- Các tính năng cơ bản cần có của một website hỏi đáp (tạo câu hỏi, trả lời câu hỏi, đánh giá tương tác với câu hỏi và câu trả lời, quản lý,…)

- Nghiên cứu về recommendation system và xử lý ngôn ngữ tự nhiên trong machine learning để áp dụng cho tính năng gợi ý câu hỏi tương tự đã có cho người dùng.

- Nghiên cứu xây dựng hệ thống chat real time

- Deploy hệ thống trên Docker.

**b. NỘI DUNG**

Giới thiệu về Natural Language Processing

Tìm hiểu Word2Vec để mô hình hoá từ ngữ thành không gian vecto

Training dữ liệu:

+ Thu thập dữ liệu từ Google Big Query dataset

+ Training với tập dữ liệu thu được

Xây dựng API

+ Tìm hiểu Flask Framework trong Python.

+ Xây dựng API cho việc tương tác gợi ý câu hỏi

+ Kết hợp với MongoDB để thực hiện việc training dữ liệu khi phát sinh dữ liệu mới

Thực hiện chức năng gợi ý câu hỏi trên Laravel (Gọi API của Flask).

Xây dựng chức năng chat real time:

Thêm tính năng tìm kiếm nâng cao

Thêm tính năng bảng xếp hạng

1. **KHỐI LƯỢNG CÔNG VIỆC ĐÃ THỰC HIỆN**

- Tìm hiểu kiến thức liên quan

- Thu thập dataset từ Google BigQuery

- Tiền xử lý dữ liệu.

- Training tập dữ liệu

- Xây dựng API sử dụng Flask Framework và test API sử dụng Postman

- Tìm hiểu và deploy được project lên docker.

**3. KHÓ KHĂN**

Cách để hệ thống có thể tự động training dữ liệu mới khi có dữ liệu mới được thêm vào MongoDB

*TP. Hồ Chí Minh, ngày 01 tháng 05 năm 2020*

**GIẢNG VIÊN HƯỚNG DẪN SINH VIÊN THỰC HIỆN**

TS. Huỳnh Xuân Phụng Trần Tấn Quý Phạm Hoàng Việt