**Question and Ansering :**

Mô tả bài toán :

Xây dựng hệ thông trả lơi hệ thông tự động dựa vào cách Xác Định Intend (dự định) người dùng . Nghĩa là close domain,intend có giới hạn cụ thể

Cần phải xác định intend là gi --> Classification

Xây dựng các câu hỏi có cùng intend :

( x(i) , y(1) ) , (x(i+1) , y(1) ) : x(i+1) , x(i) là câu diễn dịch có cung intend la y(1)

Dùng các thuật toán phân lớp để huận luyện mô hình .

y(i) = { y : 1->k} la các intend đã định nghĩa trước đó.

Flow :

Dữ liệu huấn luyện

Tiền xử lý dữ liệu

Trích xuất đặc trưng

Huần luyện mô hình

Mô hình phân lớp

Kết Quả

Câu hỏi

1. Tiền xử lý dữ liệu

* Tách từ (Token)
* Xử lý viết tắt
* Xử lý sai chính tả
* POS tagging

1. Trích xuất đặc trưng

* Bag of word
* Tf- idf

1. Thuật toán huấn luyện mô hình
2. Đánh giá mô hinh