Nghiên cứu các phương pháp phân loại văn bản

1. Nội dung nghiên cứu

* Nghiên cứu phương pháp máy học cơ bản Naive Bayes và phương pháp SVM để phân lớp các bài báo tin tức.

1. Chi tiết về phương pháp

* Nghiên cứu và cài đặt phương pháp Naive Bayes và phương pháp SVM trên bộ dữ liệu cho bài toán phân lớp văn bản bao gồm các bước sau đây:
  + Thu thập  và đọc dữ liệu
  + Tiền xử lý dữ liệu
  + Rút trích đặc trưng
  + Biểu diễn các đặc trưng thành các vector để đưa vào mô hình máy học Naive Bayes
  + Sử dụng các độ đo để đánh giá kết quả của các mô hình

Mô hình dự kiến của phương pháp Naive Bayes, Support Vector Machine: 