


# **Phân loại bình luận của trang thương mại điện tử bằng các mô hình máy học**

Danh sách nhóm:

- Trịnh Hưng Long – 18521060
- Đoàn Xuân Minh – 18521093
- Lữ Đình Phương – 18521274



GVHD: Lê Đình Duy  
Lớp: CS519.L11



# Giới thiệu đề tài

- Mua sắm online đang ngày càng phát triển.
- Nhược điểm: người dùng không thể tận mắt đánh giá chất lượng sản phẩm

➔ Hệ thống phân loại các bình luận để đánh giá chất lượng sản phẩm.  
Input: Một bình luận về một sản phẩm của khách hàng tại các trang thương mại điện tử Việt Nam.  
Output: Phân loại bình luận đó là tích cực, tiêu cực hay trung tính.



# Prepare Data

The Tiki logo is displayed in blue text within a white rounded rectangle with an orange border. The letters 'i' in 'Tiki' have blue circular accents above them.The Shopee logo features a white shopping bag icon with a red 'S' inside, followed by the word 'Shopee' in white text, all on a red rounded rectangle with an orange border.The Laz logo consists of a 3D isometric cube with orange and pink faces, and the word 'Laz' in white text, set against a dark blue rounded rectangle with an orange border.

# Data Preprocessing



## Làm sạch dữ liệu

Loại bỏ noise trong data



## Chuẩn hóa từ

Đưa dữ liệu về cấu trúc cơ bản



## Loại bỏ các stopwords

Loại bỏ các từ không cần thiết trong dữ liệu



## Vector hóa dữ liệu

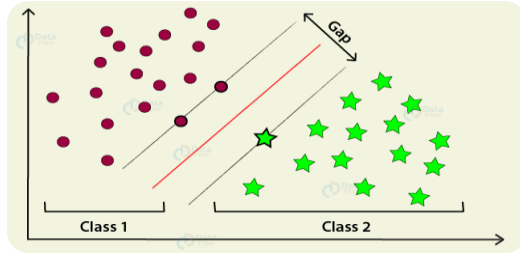
Chuyển dữ liệu về dạng vector (TF-IDF)



# Model

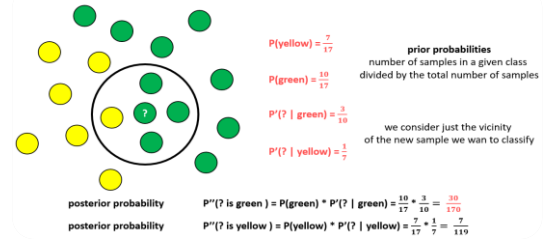
1

## SVM



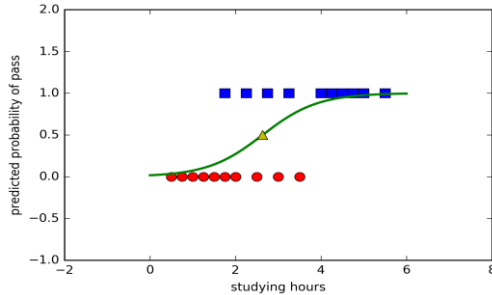
2

## Naïve Bayes



3

## Logistic Regression



4

## Random Forest

