ĐẠI HỌC THÁI NGUYÊN
TRƯỜNG ĐẠI HỌC CÔNG NGHỆ THÔNG TIN VÀ TRUYỀN THÔNG

TRẦN THỊ THANH

ỨNG DỤNG GIẢI THUẬT DI TRUYỀN VÀO PHÂN LOẠI TÀI LIỆU DẠNG VĂN BẢN

LUẬN VĂN THẠC SĨ: KHOA HỌC MÁY TÍNH

Thái Nguyên - 2012

LỜI CAM ĐOAN

Tôi xin cam đoan luận văn "**Ứng dụng giải thuật di truyền vào phân loại tài liệu dạng văn bản**" là công trình nghiên cứu của riêng tôi dưới sự hướng dẫn của **PGS.TS. Bùi Thế Hồng**. Toàn bộ phần mềm do chính tôi lập trình và kiểm thử. Tôi xin chịu trách nhiệm về lời cam đoan của mình.

Các số liệu và thông tin sử dụng trong luận văn này hoàn toàn là trung thực.

Tác giả

Trần Thị Thanh

MỤC LỤC

DANH MỤC CÁC KÝ HIỆU, CÁC TỪ VIẾT TẮT

Các từ viết tắt	Nghĩa tiếng anh	Nghĩa tiếng việt
KDD	Knowledge Discovery and Data Mining	Kỹ thuật phát hiện tri thức và khai phá dữ liệu
VSM	Vector Space Model	Mô hình không gian vector
VC	Vapnik-Chervonenkis	Kích thước VC
SVM	Support Vector Machine	Bộ phân loại Vector hỗ trợ
RBF	Radial Basis Functions	Bộ phân loại chức năng
SMO	Sequential Minimal Optimization	Tối ưu hóa tuần tự cực tiểu
TF	term frequency	Tần suất từ
k-NN	k-Nearest Neighbor	Thuật toán k-NN
WFST	Weighted Finite State Transducer	Mô hình WFST kết hợp mạng Noron
SW	Stop Words	Loại từ dừng

DANH MỤC CÁC BẢNG

DANH MỤC CÁC HÌNH VẼ