

듀얼 SMS 스팸 필터링: 그래프 기반 자질 가중치 기법

황재원^o, 고영중
동아대학교, 컴퓨터공학과
{sftcap, youngjoong.ko}@gmail.com

Dual SMS SPAM Filtering: A Graph-based Feature Weighting Method

Jae-Won Hwang^o, Young-Joong Ko
Donga University, Department of Computer Engineering

요 약

본 논문에서는 최근 급속히 증가하여 사회적 이슈가 되고 있는 SMS 스팸 필터링을 위한 듀얼 SMS 스팸 필터링 기법을 제안한다. 지속적으로 증가하고 새롭게 변형되는 SMS 문자 필터링을 위해서는 패턴 및 스팸 단어 사전을 통한 필터링은 많은 수작업을 요구하여 부적합하다. 그리하여 기계 학습을 이용한 자동화 시스템 구축이 요구되고 있으며, 효과적인 기계 학습을 위해서는 자질 선택과 자질의 가중치 책정 방법이 중요하다. 하지만 SMS 문자 특성상 문장들이 짧기 때문에 출현하는 자질의 수가 적어 분류의 어려움을 겪게 된다. 이 같은 문제를 개선하기 위하여 본 논문에서는 슬라이딩 윈도우 기반 N-gram 확장을 통해 자질을 확장하고, 확장된 자질로 그래프를 구축하여 얇은 구조적 특징을 표현한다. 학습 데이터에 출현한 N-gram 자질을 정점(Vertex)으로, 자질의 출현 빈도를 그래프의 간선(Edge)의 가중치로 설정하여 행(HAM)과 스팸(SPAM) 그래프를 각각 구성한다. 이렇게 구성된 그래프를 바탕으로 노드의 중요도와 간선의 가중치를 활용하여 최종적인 자질의 가중치를 결정한다. 입력 문자가 도착하면 스팸과 햄의 그래프를 각각 이용하여 입력 문자의 2개의 자질 벡터(Vector)를 생성한다. 생성된 자질 벡터를 지지 벡터 기계(Support Vector Machine)를 이용하여 각 SVM 확률 값(Probability Score)을 얻어 스팸 여부를 결정한다. 3가지의 실험환경에서 바이그램 자질과 이진 가중치를 사용한 기본 시스템보다 F1-Score의 약 최대 2.7%, 최소 0.5%까지 향상되었으며, 결과적으로 평균 약 1.35%의 성능 향상을 얻을 수 있었다.

주제어: 듀얼 스팸 필터링, 지지 벡터 기계, 자질 가중치, 그래프

1. 서론

SMS 스팸이란 광고나 이윤을 목적으로 다수 수신인에게 무더기로 송신된 메시지를 의미한다. 스팸은 대부분의 경우에 수신인이 원하지도 않고 관심도 없는 메시지이다. 2000년대까지는 주로 E-mail에서 이러한 스팸 메일들이 기승을 부렸으나, 현재 네트워크 인프라의 빠른 성장과 폭발적인 모바일 장비 사용자의 증가로 인하여 모바일 사용자들에게도 많은 스팸 메시지가 전송되고 있다. 이로 인해 불쾌감을 유발하고 불편을 가중시키고 있어 모바일 스팸이 사회적 이슈가 되고 있다. 현재 각 통신사들이 대량 스팸 발송 사업자 계약해지, 1일 메시지 발송량 제한 등의 노력을 하고 있으나 2012년 하반기 기준으로 1인당 1일 스팸 수신량이 0.22통에 이르고 있다 [1].

이러한 모바일 스팸을 자동으로 구별하기 위한 방법으로는 E-mail 스팸 분류를 위한 기법들이 사용되어져 왔다. 대표적인 초기 스팸 차단 방법은 목록 기반 분류 방법과 규칙기반 분류 방법 등이 있다[2]. 목록 기반 분류 방법은 대량 스팸 메시지 발신자의 블랙리스트를 만들어 이 발신자가 발송하는 메시지들을 차단하는 방법이다. 규칙기반 분류 방법은 스팸 사전에 등록된 단어들의 조합에 의한 차단 규칙을 만들고 메시지에 등장하는 단어

들이 차단 규칙에 해당되는 경우 차단하는 방법이다. 이 방법들은 블랙리스트 또는 스팸 단어 사전을 수작업으로 구축해야 한다는 단점이 있으며, 자소 분리, 특수문자 등 최근 변형된 스팸 메시지에 대한 대응이 어렵다는 단점이 있다.

이러한 문제점을 해결하기 위해 기계 학습을 통한 스팸을 차단할 수 있는 방법에 대한 연구가 활발히 진행되고 있으며, 가장 대표적인 방법은 자질을 활용한 벡터 공간 모델(Vector Space Model)이다. 그러나 SMS 메시지의 경우 한 번에 전송 가능한 글자 수의 제한으로 인해 그 내용이 짧다는 특징이 있다. 이 같은 특징은 학습 과정에서 선택된 자질만을 사용하여 스팸 여부를 판단하는 기계 학습을 통한 분류 작업을 어렵게 만들게 된다[3]. 이 같은 문제를 해결을 위하여 n-gram 방법을 통한 자질의 확장과 SMS 메시지의 구조적인 정보를 자질로 활용하는 연구도 진행되었다[4][5]. 또한, 기계 학습 기법에서는 자질의 가중치 산정 방법이 분류의 성능을 크게 좌우하게 된다. 특히 문서보다 매우 짧은 SMS 메시지의 경우 짧은 메시지에 적합한 가중치 산정 방법을 결정하는 것이 중요하다[6].

본 논문에서는 한국어 SMS 스팸 필터링을 위한 듀얼 SMS 스팸 필터링 시스템을 제안한다. 먼저, 효과적인 자질의 가중치 결정 방법과 N에 따른 N-gram 자질의 확장 방법을 비교하고, 이를 활용하여 위에서 언급한 문제를 해결하기 위하여 슬라이딩 윈도우 기반 N-gram 자질 확장 기법을 통해 자질을 확장하여 데이터 희소성 문제를 완화한다. 그리고 확장된 자질을 그래프의 정점

* 본 연구는 미래창조과학부 및 정보통신기술진흥센터의 정보통신산업 융합원천기술개발사업의 일환으로 수행하였음. [10041678, 다중영역 정보서비스를 위한 대화형 개인 비서 소프트웨어 원천 기술 개발]

(Vertex), 출현 빈도를 간선(Edge)의 가중치로 설정하여, 학습데이터를 통해 햄(HAM)과 스팸(SPAM) 그래프로 각각 표현한다. 이렇게 구성된 그래프는 정점의 중요도와 각 정점에 연결된 간선의 가중치를 통해 자질의 가중치를 결정한다. 입력 문자가 도착하면 시스템은 햄과 스팸의 그래프를 통해 2개의 자질 벡터(Feature Vector)를 생성하고 지지 벡터 기계를 통해 얻은 각 자질 벡터의 확률 값(Probability Score)의 합과 임계치(Threshold)와 비교하여 최종적인 결정을 내리게 된다. 이와 같이 임계치를 통해 특정 클래스의 정확률(Precision)과 재현율(Recall)을 조절할 수 있을 경우, False Alarm의 비율을 제어 할 수 있기 때문에 큰 장점을 제공한다. 제안하는 듀얼 SMS 스팸 필터링 시스템 구성도는 그림 1과 같다.

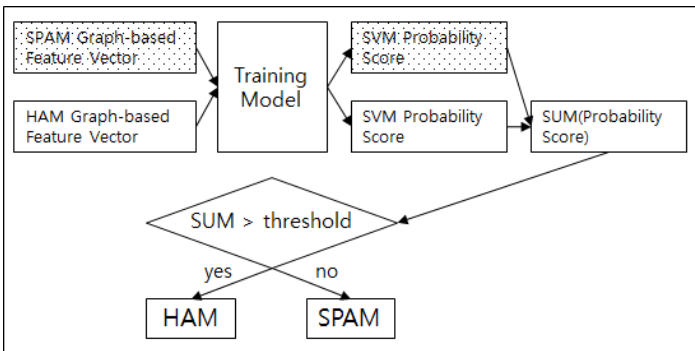


그림 1 듀얼 스팸 필터링 시스템

본 논문의 2장에서는 관련 연구를 소개하고, 3장에서는 제안하는 방법에 따른 그래프 구성 방법 및 가중치 결정 방법, 듀얼 예측 방법에 대해 기술한다. 4장에서는 자질의 가중치 결정 방법, N-gram에 따른 성능을 비교하고, 제안한 방법의 성능과 비교 평가한다. 마지막으로 5장에서는 결론과 향후 연구에 대해 제시한다.

2. 관련 연구

기계학습 기반의 스팸 차단에 관한 연구는 주로 효과적인 자질 추출 방법에 관한 연구에 초점이 맞추어져 왔다. 한국어에 대한 스팸 분류에서 효과적인 자질 선택 방법에 관한 연구는 [7][8] 등의 연구가 있으며, 이 연구에서는 신경망(Neural Network)과 Naive bayesian, 지지 벡터 기계를 사용하여 적합한 자질 추출 방법을 찾았다.

SMS 메시지의 특성상 짧은 내용으로 인해 분류를 위한 자질 부족의 극복에 초점을 맞춘 연구[4][5]도 있었다. [4]에서는 자질집합을 문자 Bigram, 문자 Trigram, 직교 최소 단어 Bigram으로 확장하여 자질 집합을 비교 실험하였다. [5]에서는 대문자의 비율, 메시지의 길이 등의 구조적인 자질을 분류를 위한 자질로 사용하였다.

문자 메시지와 같이 짧은 문장의 분류를 위해서는 구조적인 특징을 활용하는 것이 필요하다. 문자 메시지 특성상 언어 파괴가 많고 특수 문자가 많이 포함되어 있는 문장인 경우 구문 구조 분석을 통한 좋은 결과를 기대하

기 힘들다. 그래서 [9]에서는 문자 메시지에 포함된 자질을 Graph로 표현하고 각 클래스에 출현한 자질의 확률 분포를 KL-divergence를 통해 분류를 수행하였다.

3. 듀얼 SMS 스팸 필터링

3.1 슬라이딩 윈도우 기반 N-gram

짧은 문자 메시지 특성상 문장 내 출현 자질이 부족한 경우가 많다. 이 같은 현상은 학습데이터내의 자질 단어만 사용하는 기계 학습 방법에서 분류의 어려움을 겪게 만드는 원인이 된다. 그래서 본 논문에서는 슬라이딩 윈도우 기반 N-gram 확장 방식을 통해 자질을 확장하였다. 계속 새롭게 변형되는 스팸 문자의 경우 새롭게 등장하는 단어가 많기 때문에 본 연구에서는 바이그램만으로 부족하다고 판단하여 슬라이딩 윈도우에 기반하여 자질을 확장하였다. N = 3이라고 할 때 확장된 결과는 그림 2와 같다.

5555 운전 고객님	운전을 대신해 드립니다
생성된 자질	
5555 운전	
5555 고객님	
운전 고객님	
운전 운전을	
고객님 운전을	
고객님 대신해	
운전을 대신해	
운전을 드립니다	
대신해 드립니다	

그림 2 슬라이딩 윈도우 기반 N-gram 확장(N = 3)

3.2 듀얼 그래프 생성

이렇게 확장된 자질들을 바탕으로 햄과 스팸의 각 자질 그래프를 생성한다. 바이그램 자질은 두 개의 어절로 이루어져 있다. 이는 두 개의 정점이 되게 되며 이 두 정점은 하나의 간선으로 표현되게 된다. 그림 2의 예를 통해 생성된 그래프는 그림 3과 같다.

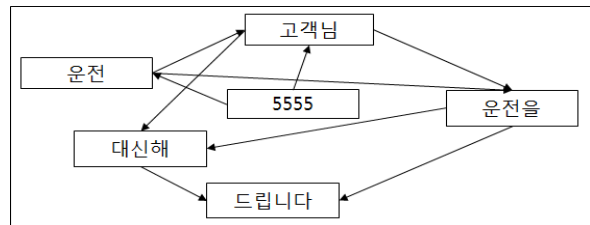


그림 3 생성 그래프

3.3 그래프 기반 자질 가중치 결정

햄과 스팸의 듀얼 그래프들은 각 클래스에서 출현한 자질들로 독립적으로 생성되게 되며, 자질의 출현 빈도에 의해서 정점과 간선의 가중치가 결정되게 된다. 이렇게 결정된 가중치를 통하여 수식 (1)을 통해 자질의 가중치를 결정한다.

$$\frac{Edge(v_i, v_j)}{avg-outgoing-edge} * important(v_j) \quad (1)$$

$$important(v_i) = outgoingEdge(v_i) + incomingEdge(v_i) \quad (2)$$

v_i, v_j 는 바이그램 자질의 각 유니그램을 의미하며 정점 $i(v_i)$ 는 정점 $j(v_j)$ 보다 선행하는 단어를 의미한다. $avg-outgoing-edge$ 는 그래프내의 노드별 평균 나가는 간선의 수를 의미하고, $Edge(v_i, v_j)$ 는 정점 i 에서 정점 j 로 나가는 간선의 수를 의미한다. $important(v_j)$ 는 정점 j 의 중요도를 의미한다. 각 정점의 중요도는 그 정점에서 나가는 간선과 들어오는 간선의 합으로 결정한다.

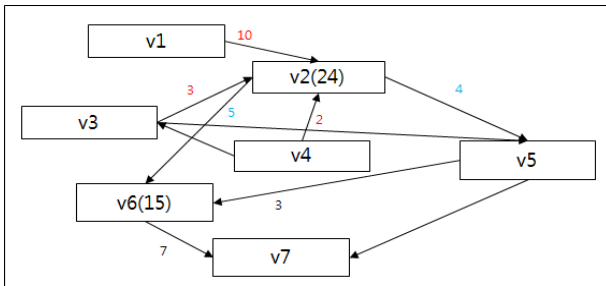


그림 4 정점과 간선의 가중치 표현

예를 들어 바이그램(대리 운전) $v2$ (대리), $v6$ (운전)의 가중치를 결정할 경우, 평균 간선의 수가 20이라고 했을 때, $5/20 * 15 = 3.75$ 의 가중치를 가지게 된다. 이렇게 결정된 가중치는 기본 이진(Binary) 가중치에 더해지게 되며, 만약 0보다 작은 값을 가지게 되면 이진 가중치로 설정되게 된다. 한국어에서 일반적으로 뒤에 오는 단어가 중요한 경우가 많기 때문에 후행 단어의 중요도만을 반영하였다. 하지만 실험의 4.5장에서 선행 단어와 선행 및 후행 단어의 중요도를 반영하여 비교 실험을 수행하였다.

유니그램의 자질인 경우 그래프로 표현하지 않고 이진 가중치를 부여하게 된다.

3.4 듀얼 예측 단계

3.2와 3.3을 통해 생성된 그래프와 자질의 가중치를 통해 입력 문자가 들어오면 햄과 스팸 기반의 2개의 자질 벡터를 생성한다. 햄 그래프 기반 자질 벡터를 클래스가 햄인 학습데이터들만 이용하여 생성한 그래프를 통해 결정된 자질의 가중치를 사용한다. 스팸 그래프 기반 자질 벡터는 스팸 학습데이터로 생성된 그래프를 이용한다. 이렇게 2개의 자질 벡터가 생성되면 학습된 지지 벡터 기계 모델로 각 자질 벡터의 SVM 확률 값(Probability Score)을 얻을 수 있다. 그림 1과 같이 얻어진 확률 값을 합하여 임계치보다 크면 햄으로 그렇지 않다면 스팸으로 결정하게 된다(햄인 경우, 확률 값이 양수가 나오도록 클래스를 부여하였다).

4. 실험 결과

본 논문의 타당성을 증명하기 위해 표 1과 같은 조건 하에 비교 실험을 진행하였다.

표 1 비교 실험 정보

항목	내용
실험 데이터	SMS 문자 데이터 (총 17,323건, 햄 9,997건, 스팸 7316건)
학습 알고리즘	Support Vector Machine (libsvm 사용)[10]
비교 대상	N-gram, Weighting Scheme
실험 방법	5-fold Cross Validation
평가 척도	F1-score, FP-rate, TP-rate

본 논문에서는 학습은 햄과 스팸의 비율을 약 8:7로 학습하였고, 테스트 시에는 실제 환경과 유사하게 테스트를 진행 위하여 테스트 데이터의 햄과 스팸의 비율을 10:1, 10:2, 10:3으로 달리하여 실험 환경을 구성하였다(1인당 1일 평균 스팸 수신량 0.22).

4.1 평가 척도

본 논문에서 사용한 TP-rate와 FP-rate는 수식 (3), (4)와 같다.

$$TPR = \frac{TP}{TP+FN} \quad (3)$$

$$FPR = \frac{FP}{FP+TN} \quad (4)$$

TP는 스팸 문자를 스팸으로 제대로 분류한 개수이고, FN은 햄 문자를 스팸으로 잘못 분류한 개수이다. TN은 햄 문자를 햄으로 제대로 분류한 개수이고, FP는 스팸 문자를 햄으로 잘못 분류한 개수이다. TPR의 경우 높을수록 많은 스팸 메시지를 잘 분류하는 것이며, FPR의 경우 햄을 스팸으로 잘못 분류한 비율이기 때문에 낮을수록 좋은 시스템이다. 현실적인 환경에서 TPR보다는 FPR이 더 중요하다고 볼 수 있으나, 이 두 개의 값은 서로 상충관계(trade-off)가 있으며, 정책에 따라 결정될 수 있다.

4.2 전처리

문자 메시지의 특성상 한글 파괴 현상이나 특수 문자가 많기 때문에 전처리 작업을 수행하였다. 표 2는 전처리의 예이다.

표 2 전처리 예

실제 메시지	수정 메시지
ㄱr'능	가능
e벤트	이벤트

모든 전처리 과정은 4.3 ~ 4.6 실험에 사용된 데이터에 적용하여 실험을 수행하였다.

4.3 N-gram, Binary vs TFIDF Scheme 비교 실험

문장의 길이가 짧은 문자 메시지에서 기존의 정보 검색에서 사용하는 방법 중 어떤 것이 적합한지를 알아보기 위하여 실험을 수행하였다. 비교 대상은 N-gram에서 Unigram과 Bigram, Weighting Scheme에서 TFIDF와 Binary의 비교 실험을 수행하였다. 햄과 스팸의 비율이 10:3인 환경에서의 성능 결과는 표 3과 같다.

표 3 기본 성능 비교 실험 결과

	TFIDF	Binary
Unigram	0.94175	0.95395
Bigram	0.94225	0.95506

실험 결과, Bigram 방식과 Binary Weighting Scheme이 우수하다는 것을 알 수 있다. 이후의 실험은 모두 Binary Weighting Scheme에 기반하게 된다.

4.4 슬라이딩 윈도우 N-gram 비교 실험

N-gram 확장을 위하여 적합한 N을 결정하는 실험을 수행 하였다. 햄과 스팸의 비율이 10:3인 환경에서의 성능 결과는 표 4와 같다.

표 4 슬라이딩 윈도우 N 결정 실험 결과

Size N	F1-score
2	0.94858
3	0.95108
4	0.94951

윈도우 사이즈가 3일 때 가장 높은 성능을 보였다.

4.5 그래프 기반 자질 가중치 결정 방법 비교 실험

수식 (1)에서 *important*에 나가는 정점뿐만 아니라 아래의 수식 (5), (6)의 경우에도 비교 실험을 수행하였다.

$$\frac{Edge(v_i, v_j)}{avg-outgoing-edge} * important(v_i) \quad (5)$$

$$\frac{Edge(v_i, v_j)}{avg-outgoing-edge} * (important(v_i) + important(v_j)) \quad (6)$$

윈도우 사이즈 3인 경우, 실험 결과는 표 5와 같다.

표 5 자질 가중치 결정 방법 실험 결과

방법	F1-Score
수식 (1)	0.95108
수식 (5)	0.95088
수식 (6)	0.95087

한국어에서 일반적으로 알려진 “뒤에 오는 단어가 중요하다” 라는 직관과 같은 결과를 얻을 수 있었다.

4.6 듀얼 예측 비교 실험

듀얼 예측 실험에서는 임계치가 중요한 역할을 한다. 이 임계치를 통하여 TP-rate와 FP-rate를 조절할 수 있기 때문이다. 적절한 임계치를 결정하기 위하여 [-2,0] 사이의 값을 0.1단위로 비교 실험을 수행하였다. 총 3가지 환경(햄:스팸 비율, 10:1, 10:2, 10:3)에서 실험을 수행하였으며, 임계치에 따른 ROC 그래프[11]의 실험 결과는 부록의 그림 5 ~ 7과 같다.

각 3가지의 환경에서 임계치(t)에 따른 가장 높은 성능은 표 6 ~ 9과 같다.

표 6 10:3 환경에서의 성능 비교

N-gram	Bigram	Window (Size = 3)	비고
Weighting	Binary	Graph (t = -0.7)	
Precision	0.94771	0.96158	+0.01388
Recall	0.96303	0.95866	-0.00436
F1-score	0.95506	0.96010	+0.00504
FR-rate	0.02861	0.01701	-0.01160

표 7 10:2 환경에서의 성능 비교

N-gram	Bigram	Window (Size = 3)	비고
Weighting	Binary	Graph (t = -0.8)	
Precision	0.93056	0.95495	+0.02439
Recall	0.96395	0.95475	-0.00920
F1-score	0.94617	0.95484	+0.00867
FR-rate	0.02861	0.01500	-0.01360

표 8 10:1 환경에서의 성능 비교

N-gram	Bigram	Window (Size = 3)	비고
Weighting	Binary	Graph (t = -1.4)	
Precision	0.88357	0.95334	+0.06978
Recall	0.96670	0.94050	-0.02620
F1-score	0.91977	0.94670	+0.02692
FR-rate	0.02861	0.00800	-0.02061

3가지 환경에서 실험 결과, 10:3 환경에서 F1-score의 최소 성능 향상인 약 0.5%의 성능 향상을 얻었고, 10:2 환경에서는 약 0.87%, 10:1 환경에서는 최고 향상 수치인 약 2.7%의 성능향상을 얻을 수 있었다. 결과적으로 평균 약 1.35%의 성능 향상을 얻을 수 있었다. 스팸 메시지가 햄에 비해 적게 들어오는 환경일수록 임계치(t)의 값이 작아질수록 성능이 향상되었으며, 3가지의 환경에서 모두 F1-Score 성능이 향상되는 결과를 얻을 수 있었다. 그리고 스팸 필터링에서 중요한 평가 항목인 FP-rate도 많은 TP-rate의 손실 없이 향상 시킬 수 있었다.

FP-rate와 TP-rate는 상충관계가 있고 정책에 따른 두 요소의 중요도가 다르기 때문에, FP-rate와 TP-rate가 모두 감소하거나(FPR이 좋아진 경우), FP-rate와 TP-rate가 모두 증가했다고(TPR이 좋아진 경우) 하여 성능이 향상되었다고 판단하기 어렵다. 그러나 F1-score에 포함된 스팸 클래스의 Recall이 TP-rate를 나타내기 때문에 FP-rate가 감소한 상황에서 F1-score의 향상은 유효한 성능 개선이라고 말할 수 있다. 성능 향상의 유효성 증명을 위하여 t-test[12]를 수행하였다. 그 결과는 표 9와 같다.

표 9 F1-score의 성능 유효성 평가

환경(햄 : 스팸 비율)	p value
10:1 환경	0.00226
10:2 환경	0.01553
10:3 환경	0.00090

3가지 환경에서의 성능에 대한 유효성 평가 결과, 모두 $p < 0.05$ 의 값으로 유효하다는 결과를 얻었다.

5. 결론

본 논문에서는 한국어 SMS 스팸 필터링을 위하여 듀얼 스팸 필터링 시스템을 제안하였다. 짧은 문장인 경우 기계 학습 기법 적용시에 발생할 수 있는 데이터 희소성 문제(Data Sparseness Problem)를 완화하기 위하여 슬라이딩 윈도우 기반 N-gram 방법을 통하여 자질을 확장하였으며, 확장된 자질을 그래프로 구성하여, 스팸과 햄의 독립적인 그래프를 생성하였다. 이렇게 생성된 2개의 그래프를 통하여 입력 문자 메시지가 들어오면 햄과 스팸 그래프에 따른 자질의 가중치를 가진 2개의 자질 벡터를 생성하고 지지 벡터 기계를 통해 분류하여 얻어진 확률 값을 바탕으로 스팸 여부를 결정하였다. 실험 결과 일반적인 바이그램과 한국어 스팸 필터링에 유용하다고 알려진 이진 가중치를 설정한 방법에 비해 3가지의 실험 환경에서 평균 F1-score의 약 1.35%의 성능 향상을 얻을 수 있었다.

향후 연구로는 구조적 특성을 더 반영할 수 있는 자질 확장 방법과, 그래프를 통한 효과적인 가중치 측정 방법에 관한 연구를 수행할 것이다. 그리고 규칙기반의 분류기와 기계 학습 분류기의 하이브리드 방식의 분류에 관한 연구를 진행할 것이다.

참고문헌

[1] 방송통신위원회, 한국인터넷진흥원 보고서, "2012년 하반기 스팸 유통현황 측정 결과", 2013.
 [2] 이신영, 길아라, 김명원, "링크구조분석을 이용한 스팸 메일 분류", 정보과학회논문지:소프트웨어 및 응용, 제34권, 제1호, pp. 30-39, 2007.
 [3] J. M. Gomez et al., "Content Based SMS Spam Filtering", In Proceedings of the 2006 ACM Symposium on Document Engineering, pp. 107-114, 2006.

[4] G. V. Cormak et al., "Spam Filtering for Short Message", In Proceedings of the ACM Sixteenth Conference on Information and Knowledge Management, pp. 313-320, 2007.
 [5] A. K. Uysal et al., "The Impact of Feature Extraction and Selection on SMS Spam Filtering", In Electronics & Electrical Engineering, Vol. 19, No.5, pp. 67-72, 2013.
 [6] 이성욱, "카이제곱 통계량과 지지 벡터 기계를 이용한 스팸 메일 필터", 한국정보처리학회지, 제17-B권, 제132호, pp. 249-254, 2010.
 [7] 홍진혁, 류중원, 조성배, "실세계의 FAQ 메일 자동 분류를 위한 문서 특징추출 방법의 성능비교", 한국정보과학회 봄 학술 발표논문집, 제28권, 제1호, pp. 232-234, 2001.
 [8] 송철환, 유성준, "문서 분류 알고리즘을 이용한 한국어 스팸 문서 분류 성능 비교", 한국정보과학회 가을 학술 발표논문집, 제33권, 제2호, pp. 222-225, 2006.
 [9] M. Z. Rafique and M. Abulaish, "Graph-based Learning Model for Detection of SMS Spam on Smart Phones", In Proceedings of 8th International Wireless Communications and Mobile Computing Conference (IWCMC), pp. 1046-1051, 2012.
 [10] C. C. Chang and C. J. Lin, "LIBSVM : a library for support vector machines", In ACM Transactions on Intelligent Systems and Technology, Vol.2, pp. 27:1-27:27, 2011.
 [11] T. Fawcett, "ROC graphs: Notes and Practical Considerations for Researchers", Machine Learning, Vol.31, 2004.
 [12] J. Demsar, "Statistical comparisons of classifiers over multiple data sets", The Journal of Machine Learning Research, Vol.7, pp.1-30, 2006.