ISSN : 1226-7244 (Print)
ISSN : 2288-243X (Online)
논문번호 19-04-40
278
j.inst.Korean.electr.electron.eng.Vol.23,No.4,1393~1399,December 2019
http://dx.doi.org/10.7471/ikeee.2019.23.4.1393

Lifelong Machine Learning 기반 스팸 메시지 필터링 방법

A Method for Spam Message Filtering Based on Lifelong Machine Learning

안 연 선*, 정 옥 란*★

Yeon-Sun Ahn*, Ok-Ran Jeong*★

Abstract

With the rapid growth of the Internet, millions of indiscriminate advertising SMS are sent every day because of the convenience of sending and receiving data. Although we still use methods to block spam words manually, we have been actively researching how to filter spam in a various ways as machine learning emerged. However, spam words and patterns are constantly changing to avoid being filtered, so existing machine learning mechanisms cannot detect or adapt to new words and patterns. Recently, the concept of Lifelong Learning emerged to overcome these limitations, using existing knowledge to keep learning new knowledge continuously. In this paper, we propose a method of spam filtering system using ensemble techniques of naive bayesian which is most commonly used in document classification and LLML(Lifelong Machine Learning). We validate the performance of lifelong learning by applying the model ELLA and the Naive Bayes most commonly used in existing spam filters.

요 약

인터넷의 급속한 성장으로 데이터의 송수신의 편리성과 비용이 들지 않는다는 장점 때문에 매일 수백만 건의 무차별적인 광고성 스팸 문자와 메일이 발송되고 있다. 아직은 스팸 단어나 스팸 번호를 차단하는 방법을 주로 사용하지만, 기계 학습이 떠오름에 따라 스팸을 필터링하는 방법에 대해 다양한 방식으로 활발히 연구되고 있다. 그러나 스팸에서만 등장하는 단어나 패턴은 스팸 필터링 시스템에 의해 걸러지지 않기 위해 지속적으로 변화하고 있기 때문에, 기존 기계 학습 메커니즘으로는 새로운 단어와 패턴을 감지, 적응할 수 없다. 최근 이러한 기존 기계 학습의 한계점을 극복하기 위해 기존의 지식을 활용하여 새로운 지식을 지속적으로 학습하도록 하는 Lifelong Learning(이하 LL)의 개념이 대두되었다. 본 논문에서는 문서 분류에 가장 많이 사용되는 나이브 베이즈와 Lifelong Machine Learning(이하 LLML)의 앙상블 기법을 이용한 스팸 메시지 필터링 방법을 제안한다. 우리는 기존 스팸 필터링 시스템에 가장 많이 사용되는 나이브 베이즈와, LLML 모델 중 ELLA를 적용하여 LL의 성능을 검증한다.

Key words : Spam Filtering, Naive Bayes Classifier, Lifelong Machine Learning, ELLA

* Dept. of Software, Gachon University
★ Corresponding author
E-mail : orjeong@gachon.ac.kr, tel : +82-31-750-5831

※ Acknowledgment
This research was supported by Basic Science Research Program through the NRF(National Research Foundation of Korea), and the MSIT(Ministry of Science and ICT), Korea, under the National Program for Excellence in SW supervised by the IITP(Institute for Information & communications Technology Promotion) (Nos.NRF 2019R1A2C1008412, 2015-0-00932).
Manuscript received Dec. 6, 2019; revised Dec. 26, 2019; accepted Dec. 60, 2019.

Ⅰ. 서론

인터넷이 발달함에 따라 메시지 송수신을 위한 수단으로 이메일이 매우 활발하게 사용되었다. 현대 사회에서는 스마트폰 메신저 앱이 기하급수적으로 증가하고 있으며, 이메일 송수신을 대체하고 있다. 메신저 앱의 인스턴트 메시지는 휘발성이고, SNS는 보안이 낮기 때문이라는 지적도 있으나, 최근에는 이 둘이 개인의 프라이버시 영역이라는 인식도 적지 않다. 이 때문에 기업에서는 여전히 업무용 커뮤니케이션 수단으로써 이메일 사용이 거의 절대적이다. 하지만 송수신의 편리성과 비용이 들지 않는다는 장점으로, 많은 업체들이 무차별적인 광고성 메일을 발송하고 있다. 이를 스팸 메일이라고 불리며, 메일 뿐만 아니라 일반 우편, 게시판, 문자 메시지, 전화, SNS 쪽지 기능 등 여러 수단을 통해 수신되고 있다.

일반 사용자들은 이러한 스팸 메일이나 문자를 지우고 정리하는 데만 적지 않은 시간을 소요하고, 네트워크상에서도 엄청난 양의 패킷이 낭비되면서 스팸이 사회문제로 부각되었다. 이러한 문제를 해결하기 위해 정보통신부에서는 '정보통신망 이용촉진 및 정보보호 등에 관한 법률 시행령 및 시행규칙 개정안'을 마련해 두었으며, KISA에서는 불법스팸대응 시스템을 통해 대책을 마련하고 서비스하고 있다. 이들에 따르면, 모든 광고성 메일은 '광고'라는 단어를 필히 넣어야 하며, 스팸으로 필터링되기 쉽게 '@'를 첨가해야 한다고 규정하였지만 이를 준수하지 않은 경우가 훨씬 많다.

최근에는 기계 학습으로 데이터 분석을 통한 데이터의 특징 및 자주 나타나는 패턴을 파악할 수 있게 되면서, KISA에서는 일반 사용자들로부터 신고된 스팸 메일의 특징과 언어적 패턴을 파악하여 필터링할 수 있도록 노력을 기울이고 있다. 문서 분류에 가장 많이 사용되는 나이브 베이즈 분류기(Naive Bayesian Classifier)는 간단한 알고리즘만으로 단어의 빈도수와 단어들 사이의 출현 유무를 확률적으로 계산하여 판단함으로써 서포트 벡터 머신(Support Vector Machine)과 같이 더 발전된 방법들과도 충분한 경쟁력을 보인다[1],[2].

그러나 스팸은 이러한 필터링 시스템에 탐지되지 않기 위해 시간이 갈수록 변화한다. 이전에는 특수문자를 가득 써가며 광고했다면, 요즘은 많이 줄어 꽤 정상적인 문자 같아 보인다. 띄어쓰기를 생략하거나, 모든 음절을 띄어쓰기 하여 사람은 알아보지만 필터링 시스템이 알아볼 수 없는 문자로 표현되곤 한다. 여기서 기존 기계 학습의 문제점을 알 수 있다. 과거에 국한된 데이터 셋만을 가지고 훈련시키기 때문에, 훈련 셋에 없는 단어나 패턴은 지속적으로 발견해낼 수 없다. 빠르게 변하는 현실 세계에 대응하기 위해서는 끊임없이 지식을 학습해야 하고 새로운 개념을 배울 수 있어야 한다.

이러한 현안을 해결하기 위해, 본 논문에서는 LLML을 기반으로 스팸을 필터링하는 방법에 대해 제안한다.

Ⅱ. 관련연구

1. 스팸 필터링 시스템

스팸 필터링은 무차별적인 광고에 일반 사용자들이 불편함을 느끼기 시작한 이래로 지속적인 연구가 이뤄지고 있다. 또한 기계학습(machine learning)의 등장과 함께 지도학습(supervised-learning)의 가장 대표적인 예로 알려져 있다. 현재까지 많은 강력한 분류기들이 나왔지만 그 중 가장 간단하면서도 다른 알고리즘에 뒤지지 않는 경쟁력을 가진 다중 분포 나이브 베이즈 분류기(Multinomial Naive Bayes Classifier)가 스팸 필터링 관련 연구들에서 많은 성공을 거두었다[3],[4]. 서포트 벡터 머신과 같은, 나이브 베이즈 보다 더 강력한 분류기들이 있지만 알고리즘 자체가 복잡하여 덜 널리 사용되고 있다.

2. Lifelong Machine Learning(LLML)

LLML은 지속적으로 학습하고, 이전에 배운 지식을 축적하며, 이를 미래의 학습 및 문제 해결에 적용할 수 있도록 하는 발전된 기계 학습 패러다임이다[5].

현재 지배적인 기계 학습 패러다임은 독립적인 환경에서 사용되고 있다. 학습 데이터셋이 주어지고, 이 데이터만을 잘 표현할 수 있는 모델을 만들기 위해 기계 학습 알고리즘을 적용시킨다. 이 학습 방법은, 배운 지식을 유지하고 이를 이후의 학습에 사용하지 않는다. 물론 좋은 성능을 내지만, 이를 위해서는 꽤 많은 양의 학습을 위한 데이터셋이 필요하며, 학습으로 만들어진 모델 또한 해당

데이터셋에만 좋은 성능을 낼 뿐이다. 이를 폐쇄된 학습 환경(closed environment)라고 한다.

이와는 대조적으로 인간은 단 몇 가지의 예시로써 수많은 태스크를 개방된 학습 환경(open environment)에서 역동적이고 효과적으로 학습할 수 있다. 그 이유는 인간의 학습은 수 가지의 예시로 이루어지는 것이 아닌, 그동안 쌓아온 경험과 지식에 기반하기 때문이다. LL은 이러한 인간의 기능을 달성하는 것을 목표로 한다[6]. 챗봇(Chat-bot), 자율주행차, AI 시스템과 같은 어플리케이션들은, 이들이 잘 작동하기 위해서 지속적으로 새로운 것을 배울 수 있는 역동적이고 개방적인 환경에 대처할 필요가 있기 때문에 이러한 능력이 요구되고 있다.

본 논문에서는 LLML 기법의 ELLA(Efficient LL Algorithm) 모델을 활용하여 지속적인 학습을 하되, 스팸 유형별 모델들이 공유된 지식 기반을 유지하여 서로 다른 도메인의 지식 또한 학습할 수 있도록 한다.

III. Lifelong Machine Learning 기반 스팸 필터링 시스템

본 논문에서 제안하는 LLML 기반의 스팸 필터링 시스템은 그림 1과 같이 설계되었다. 시스템은 나이브 베이즈와 LLML 두 부분으로써 스팸을 필터링한다. 나이브 베이즈 분류기에서는 단어들이

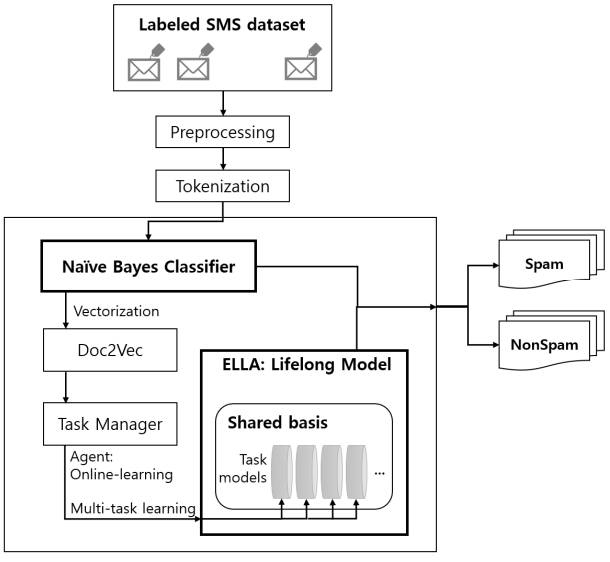

Fig. 1. LLML based spam filtering system.
그림 1. LLML 기반 스팸 필터링 시스템

스팸에서 등장할 확률, 비스팸에서 등장할 확률을 계산하여 새로 입력된 SMS가 스팸인지, 그렇다면 도박, 대리운전, 선거 등 어떤 유형에 속하는 지를 계산한다. 이 때 입력된 SMS가 스팸이고 각 스팸 유형에 해당할 확률 값을 추출했을 때 최댓값이 0.65 이하면, 즉 어느 유형에 확실하게 속한다는 뚜렷한 차이가 보이지 않는다면, LLML 모델에 재학습시킴으로써 유형을 더 정확히 분류하도록 한다.

1. 나이브 베이즈 분류기

나이브 베이즈 분류기는 조건부 확률을 계산하는 방법 중 하나이다. P(A)는 A가 일어날 확률, P(B)는 B가 일어날 확률, P(B|A)는 A가 일어났을 때 B가 일어날 확률, P(A|B)는 그 반대일 확률이다. 이 때 아래와 같은 수식 (1)로 P(B|A)로 P(A|B)를 구할 수 있다.

$$P(A|B) = P(B|A) * P(A) * P(B) \tag{1}$$

이를 스팸 필터링 시스템에 적용하여, 입력 텍스트인 SMS가 주어졌을 때 이 입력 텍스트가 스팸인지 비스팸인지 구분하기 위한 확률을 다음 수식 (2)와 같이 구할 수 있다. 스팸일 확률을 P(S): spam, 비스팸일 확률을 P(H): ham, 입력 값 SMS를 P(C): content라고 표기한다.

$$\begin{aligned} P(H|C) &= P(C|H) * P(H) / P(C) \text{ or} \\ P(S|C) &= P(C|S) * P(S) / P(C) \end{aligned} \tag{2}$$

SMS가 입력되었을 때, 비스팸일 확률이 더 크다면 정상적인 SMS로, 그렇지 않다면 스팸으로 분류한다. 이 때, 두 경우 모두 P(C)로 나누는 계산을 하고 있기 때문에 생략한다.

기본적으로 나이브 베이즈 분류기는 모든 특성(feature)이 서로 의존하지 않고 독립적이라고 가정하기 때문에, 입력 SMS에 포함하고 있는 단어들을 특성으로써 확률을 계산한다. 각 단어가 비스팸 데이터에서 나타날 확률과 스팸 데이터에서 나타날 확률을 각각 계산한 후 다음과 같이 수식 (3)으로 나타낼 수 있다.

$$P(H|C) \propto P(word_1|H) * P(word_2|H) * ... * P(H) \tag{3}$$

단어들이 독립적이고 출현 빈도수가 어떻게 되느냐에 따라 값이 달라지기 때문에, Bag-of-Words (BoW)와 같이 SMS 단어의 순서, 즉 의미는 무시하

고 오직 단어의 빈도수만을 고려하는 방식이다[7].

본 논문에서는 단순히 스팸과 비스팸을 분류하기 보다는 다중 분포 나이브 베이즈 분류기를 사용하여 스팸의 유형(class)을 가장 대표적인 도박, 선거, 대리운전, 금융사기, 그 외에 스팸(기타)과 함께 비스팸까지 총 여섯 가지로 나누어, 개별 단어가 개별 유형에 속할 확률을 계산했을 때 입력 SMS가 스팸인지, 어떤 유형에 속하는 지 분류한다.

2. Lifelong Machine Learning model: ELLA

LLML은 open environment에서의 학습을 가정하기 때문에 학습과 실험 시에 볼 수 있는 데이터의 유형이 다를 수 있음을 전제로 한다. LLML은 전이 학습(Transfer Learning)의 다중 작업 학습(Multi-task Learning)을 기반으로 학습에 사용되는 데이터의 도메인이나, 작업(task), 분포(distribution)가 다를 수 있는 환경에서, 하나의 모델이 여러 작업을 동시에 학습할 수 있도록 하며, 훈련 데이터 셋으로 구한 모델 파라미터를 이후에 입력될 새로운 데이터에 맞게 미세 조정(fine-tuning)함으로써 다양한 데이터를 표현할 수 있도록 한다[8].

본 논문에서는 LL 환경에서 온라인 다중 작업 학습을 위해 ELLA(Efficient Lifelong Learning Algorithm) 모델을 사용한다[9]. ELLA는 모든 작업에 대해 작업 모델을 개별 구축하고 이 모든 모델에 대해 희박한 공유 기반을 유지하며, 그 기반으로부터 지식을 전이하여 새로운 작업을 학습하고, 계속적으로 학습함에 따라 기반을 다듬어 모든 작업에 걸쳐 성능을 극대화한다.

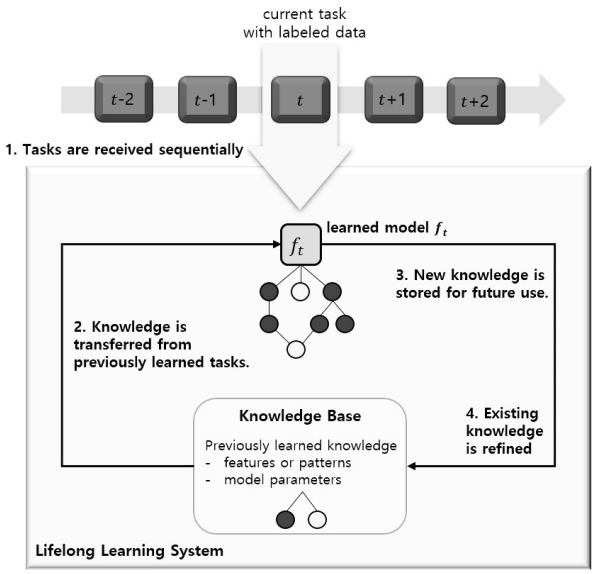

Fig. 2. Basic Lifelong learning process.
그림 2. 기본 LL 학습 과정

ELLA가 기반으로 하는 기본 LL 시스템은 그림 2와 같다. 시스템 에이전트(agent)는 라벨링된 학습 데이터(작업)를 미니 배치(mini-batch) 사이즈 만큼 계속적으로 받고, 각 작업 t에서 데이터 인스턴스 X를 정답 Y로 매핑하는 작업 모델 f를 구축한다. 입력된 데이터가 기존에 알고 있던 작업일 경우 해당 작업 모델 f를 업데이트 하고, 새로운 유형의 작업일 경우 새로운 작업 모델 f를 시스템에 추가한다. 이 때 각 작업 모델 f는 해당 작업에 특정한 파라미터 벡터인 세타에 의하며, 이 세타는 작업들 사이의 관계를 모델링하기 위해서 전체 모델 컴포넌트를 공유하는 선형 결합 모델이라고 가정한다.

$$\theta^{(t)} = Ls^{(t)} \tag{4}$$

작업 모델이 공유하는 모델 컴포넌트 L을 기반으로 가중치 벡터 s와 세타를 구하도록 한다. 이 때 가중치 벡터 s는 학습된 각 모델 컴포넌트가 최대한 재사용 가능한 지식을 포착하도록 하기 위해 희박하게 유지한다.

아래의 수식 (5)로써 공유된 구조에서 예측 손실을 최소화하도록 한다.

$$e_T(L) = 1/T\sum_{t=1}^{T}\min_{s^{(t)}} * \left\{1/n_t\sum_{i=1}^{n_t} L(f(x_i^{(t)};Ls^{(t)}), y_i^{(t)}) + \mu\|s^{(t)}\|_1\right\} + \lambda\|L\|_F^2 \tag{5}$$

작업의 데이터 인스턴스(X, y)에 대한 작업 모델을 y와 함께 손실 값 계산, 이를 L1 정규화를 거친 가중치 벡터 s와 함께 최소화 시킨다. 위 식으로써 최적 값을 찾기 위해 두 최적화 단계를 반복함으로써 값이 수렴하도록 한다.

(1) 작업의 가중치 벡터 s를 고정한 채 모델 컴포넌트를 최적화 한다.
(2) 모델 컴포넌트를 고정한 채 작업의 가중치 벡터 s를 최적화 한다.

그러나 이 경우 작업 모델 사이의 의존성이 생기

기 때문에 제2차 테일러 확장 법을 사용하여 수식 (5)를 근사화한다: 수식 (6) 또한 작업의 수가 많아질수록 s를 계속해서 다시 계산해야 하는 비용이 발생하게 되는데, 위 모델에서는 학습 중이던 작업 데이터의 제일 마지막에만 s를 업데이트 하고, 다른 작업 훈련 시 이를 업데이트 하지 않음으로써 작업의 수가 아무리 증가하여도 작업 모델의 성능에 크게 영향을 주지 않음을 보여준다.

$$g_T(L) = 1/T \sum_{t=1}^{T} \min_{s^{(t)}} * \left\{ 1/n_t \|\theta^{(t)} - Ls^{(t)}\|^2_{D(t)} + \mu \|s^{(t)}\|_1 \right\} + \lambda \|L\|^2_F \quad (6)$$

본 논문에서는 나이브 베이즈로 스팸을 빠르게 우선 분류하고, 변하는 스팸의 유형과 패턴에 적응하기 위해 LLML 모델 ELLA에 적용하여 더 정확히 분류하도록 한다. 각 작업을 스팸의 유형으로 가정하여, 각 작업 모델은 도박 스팸인지 비스팸인지, 또는 대리운전 스팸인지 비스팸인지 등을 예측할 수 있다.

Ⅳ. 실험

우리는 나이브 베이즈 분류기와 ELLA를 이용하여 스팸에 지속적으로 대응하는 스팸 필터링 시스템을 구현한다. 제안하는 시스템의 검증을 위해 나이브 베이즈 분류기만 사용했을 때와 ELLA를 적용했을 때의 정확도를 비교한다.

1. 실험 환경 및 방법

실험을 위해 KISA에서 제공받은 23,170건의 스팸 문자 데이터를 사용하였다. 데이터는 총 19개의 스팸 카테고리로 분류되어 있으나 보다 명확한 분류를 위해 신고된 빈도수 상위 5가지와 나머지 스팸, 그리고 비스팸으로 총 7가지를 분류하였다.

SMS 텍스트 데이터는 학습을 위해 Konlpy를 이용한 전처리 후 Doc2Vec 임베딩을 통해 200차원의 벡터 값으로 변환시켰으며, LLML 모델 ELLA의 개별 작업 모델은 Logistic Regression으로 학습을 진행하였다[10],[11].

위의 실험은 Google Colab GPU 환경에서 구현하였다.

Table 1. Number of spam reports by category.
표 1. 카테고리별 스팸 신고 건수

Label	Category	Number
0	Gambling	6363
1	Etc. (in spam)	4619
2	Election	2615
3	Chauffeur service	1684
4	Non Spam	1637
5	Illegal loan + Finance	2424
6	Others	3828
Total	Total	23,170

2. 실험 결과

표 2는 나이브 베이즈 분류기만 사용하여 전체 데이터를 표 1의 여섯 카테고리로 분류한 결과이다. 학습 데이터셋으로 훈련시킨 모델의 정확도는 0.917, 테스트 데이터셋으로 검증한 모델의 정확도는 0.848, 그리고 각 카테고리 분류 정확도는 평균 0.85의 f1-score를 보였다.

Table 2. Results of Naive Bayes Spam Filter.
표 2. 나이브 베이즈 스팸 필터 결과

Label	Precision	Recall	F1-score
0	0.88	0.91	0.90
1	0.76	0.79	0.78
2	0.92	0.96	0.94
3	0.98	0.95	0.96
4	0.77	0.54	0.63
5	0.82	0.89	0.85
6	0.84	0.79	0.82
Accuracy	.	.	**0.85**

표 3은 표 1의 데이터셋을 ELLA 모델에 적용시켜 카테고리 분류한 결과이다. 단, ELLA의 작업 모델 각각은 비스팸인지 아닌지를 구분하는 이진 분류이기 때문에, 표 1의 7개 유형 중 비스팸을 뺀 6개 유형을 비스팸과 비교한다. 데이터는 미니배치 사이즈만큼 모델의 입력 값으로 하되, 하나의 유형을 가진 데이터로만 구성한다. 개별 모델은 스팸 유무를 판단하는 이진 분류로 로지스틱 회귀(logistic regression)를 사용하였다. 학습 데이터셋으로 훈련시킨 작업 모델의 평균 정확도는 0.918,

테스트 데이터셋으로 검증한 정확도는 0.917, 그리고 각 카테고리 분류 정확도는 평균 0.91의 f1-score를 보였다.

표 2에서 나이브 베이즈 스팸 필터 결과는 accuracy 평균 0.85를 보여주었으나, 표 3에서 ELLA의 스팸 필터 결과는 accuracy 평균 0.91로 정확도가 개선되었다.

위 실험으로 LL을 적용한 모델이 더 잘 학습하여 더 좋은 성과를 낼 수 있었음을 확인하였다. 잡음이 유독 많은 데이터셋이었지만 그럼에도 불구하고 높은 성능을 보여주었고, 스팸뿐만 아니라 기존 기계 학습을 적용한 데이터에 대해 LL 메커니즘을 적용하면 보다 향상된 성능을 기대할 수 있을 것으로 보인다.

Table 3. Results of ELLA Spam Filter.
표 3. ELLA 스팸 필터 결과

Label	Precision	Recall	F1-score
0	0.94	1.00	0.97
1	0.93	0.99	0.96
2	0.92	0.94	0.93
3	0.95	1.00	0.97
4	0.93	0.96	0.94
5	0.93	0.99	0.96
Accuracy	.	.	**0.91**

Ⅴ. 결론

스팸은 시간이 지남에 따라 필터링 시스템에 걸러지지 않기 위해 새로운 패턴으로, 새로운 유형으로 계속 변화한다. 이에 맞춰 스팸 필터링 시스템도 끊임없이 발전하여 잘 걸러낼 수 있는 것이 중요하다.

LLML은 제한된 학습 데이터셋에 국한되지 않고 다양한 도메인에서 더 효율적으로 학습할 수 있는 매우 중요한 메커니즘이다.

여러 작업 모델을 개별적으로 구축하고 점진적으로 학습하기 때문에 스팸 유형별 특징을 각각 반영하고 기존 나이브 베이즈 분류기보다 우수한 성능을 낼 수 있었음을 실험을 통해 검증했다.

또한 작업 모델들 사이에 공유 기반을 유지하기 때문에 좀 더 다양한 유형의 스팸이 적은 수로 등장해도 좋은 분류 성과를 보여줄 것으로 기대된다.

References

[1] JM Gomez Hidalgo, GC Bringas, EP Sanz, and FC Garcia, “Content based SMS spam filtering,” *Proceedings of the 2006 ACM symposium on Document engineering*, pp.107-114, 2006. DOI: 10.1145/1166160.1166191

[2] Andrew McCallum and Kamal Nigam, “A Comparison of Event Models for Naive Bayes Text Classification,” *AAAI-98 on learning for text categorization*, 1998.

[3] Le Zhang, Jingbo Zhu, and Tianshun Yao, “An Evaluation of statistical spam filtering techniques,” *ACM Transaction on Asian Language Information Processing*, pp.243-269, 2006. DOI: 10.1145/1039621.1039625

[4] Vangelis Metsis, “Spam Filtering with Naive Bayes-Which Naive Bayes?,” *CEAS*, 2006.

[5] Zhiyuan Chen and Bing Liu, “*Lifelong Machine Learning, Second Edition,*” Morgan & Claypool publishers, 2018.

[6] Zhiyuan Chen, Nianzu Ma, and Bing Liu, “Lifelong learning for sentiment classification,” *ACL*, pp 750-756, 2015.

[7] Ion Androutsopoulos and John Koutsias, “An Evaluation of Naive Bayesian Anti-Spam Filtering,” *ECML*, pp.9-17, 2000.

[8] Abhishek Kumar and Hal Daume Ⅲ, “Learning Task Grouping and Overlap in Multi-Task Learning,” *arXiv:1206.6417*, 2012.

[9] P Ruvolo and E Eaton, “ELLA: An efficient lifelong learning algorithm,” *ICML*, 2013.

[10] Tomas Mikolov, Kai Chen, Greg Corrado, and Jeffrey Dean, “Efficient Estimation of word Representations in Vector Space,” *arXiv:1301.3781 v3*, 2013.

[11] Quoc Le, Toman Milokov, “Distributed Representations of Sentences and Documents,” *Proc. of the 31st ICML*, 2014.

BIOGRAPHY

Yeon-Sun Ahn (Member)

2020 : BS degree in Software, Gachon University.
2019~present : MS student in Software, Gachon University

Ok-Ran Jeong (Member)

2005 : PhD degree in Computer Science and Engineering, Ewha Womans University.
2006: Postdoctoral Researcher, Seoul National University
2007: Postdoctoral Researcher, Univ. of Illinois at Urbana-Champaign
2008~2009: Research Professor, Sunkyunkwan Univ.
2009~2019 : Associate Professor, Gachon University.