정 보 보 호 학 회 지
제27권 제3호, 2017. 6

AISecurity를 위한 베이지안 추론 기술 소개

윤 지 원*

요 약

최근의 인공지능과 기계학습 기술이 과학기술 전반에 걸쳐서 적용되고 있다. 정보보안 분야에서도 인공지능 기술이 다양하게 적용되어 여러 가지 우수한 성능의 제품과 기술들이 나오고 있다. 이러한 시점에 인공지능과 기계학습의 원천 이론 중 하나인 베이지안 추론 (Bayesian Inference)기술에 대한 소개를 하고자 한다. 특히, 정보보호를 연구하는 연구자들에게 베이지안 기술의 기초부터 활용에 이르는 영역을 선보이며 악성코드 분석과 함께 카드사기탐지 기술과 관련하여 베이지안 추론 기술의 적용가능성을 소개한다.

Ⅰ. 서 론

최근에 빅데이터기술과 사물인터넷 기술이 우리사회 전반에 큰 변화를 몰고 오고 있다. 특히, 빅데이터 분석 및 사물인터넷 환경에서의 제어에 가장 중요한 원천기술로 여겨지는 기술이 바로 인공지능 기술로 4차 산업혁명과 산업자동화를 넘어 정보화 사회의 가장 주축이 되었다. 특히 인공지능의 한 분파인 기계학습의 중요성은 나날이 증가되어왔으며 빅데이터라는 대용량 데이터의 출연으로 인해 기존에 불가능해보였던 다양한 문제들이 해결되고 있는 상황이다. 예를들어, 페이스 북에서의 고도의 정확도를 보이는 실시간 얼굴 인식, 동영상의 캡션을 자동으로 만들어주거나 새로운 뉴스들을 자동으로 만들어주거나 기계에 의해서 만들어지는 그림그리기와 작곡 등 인간이 담당하던 영역이 더 이상 인간만의 영역이 아니게 되었다. 이러한 변화에 중심에 바로 기계학습이라는 분야가 있어왔다. 최근의 알파고를 통해서 많은 국민들이 딥러닝과 몬테카를로 기술에 대해서 이름을 들어보기까지 하였다. 본 논문에서는 이러한 현대사회의 변화의 바람을 일으키고 있는 기계학습에서 가장 중요한 원천기술 중 하나로 여겨지고 있는 베이지안 추론 (Bayesian Inference)기술을 소개하고자 한다. 베이지안 추론의 소개와 더불어서 최근에 정보보안에서도 선보이고 있는 인공지능과 정보보안의 융복합적인 분야인 AISecurity라는 분야에서 다루어지는 대표적인 두 가지 주제인 들을 베이지안 관점에서 다루고자 한다.[1)]

Ⅱ. 배경지식

2.1. 문제풀이와 목적함수

우선 우리가 임의의 해결해야 하는 다양한 문제들을 아래와 같이 갖고 있다고 가정하자.

1) 주어진 1년치 카드 사용기록을 기반하여 현재 시간에서의 새로운 카드사용이 거짓사용자의 불법사용인지 적법한 사용자에 의한 판별이 가능한가?
2) 안드로이드 핸드폰에 홍채인식이나 지문인식을 통해 접근한 사람이 적법한 사용자인지 판단하는 생체인증(Biometrics) 기술을 이용하여 사용자 인증(Authentication) 및 인가(Authorization)가 가능한가?
3) 망관리자가 외부의 공격자로부터 내부망에 DDoS공격을 받고 있는지를 사전에 탐지하고자 하는 문제
4) 프로그램들의 동적 분석을 통한 악성코드 분석이 가능한가?
5) 주어진 파형이 있을 때 숨겨진 비밀키(Key)값은 무엇인가?

1) 본 논문은 2017년 5월에 있었던 NetSecKR에서 발표한 튜토리알의 내용의 연장선상에 있다고 볼 수 있다. 그러므로, 추가적인 내용이 필요한 독자는 signal.korea.ac.kr/ai_seminar 사이트에 들어가서 슬라이드와 실습 데이터를 얻을 수 있다.

* 고려대학교 사이버국방학과 / 정보보호대학원 교수(jiwon_yoon@korea.ac.kr)

위와 같은 문제들은 현재 정보보호 분야에서 가장 대표적인 문제로 받아들일 수 있다. 해당 문제들를 풀기 위해서 우리에게 주어진 데이터는 관찰값(observation)이라고도 불리며 이것을 우리는 y라는 것으로 표현한다. 이 값은 실수이거나 정수이거나 또 다른 형태의 데이터일 수 있다. 예를 들어 주어진 이미지를 갖고 해당 이미지에 있는 사람의 성별을 추론하는 문제의 경우 관찰 값인 y는 바로 그 이미지이다. 그리고 우리가 풀고자 하는 문제의 답안을 담을 변수가 필요한데 그것을 우리는 x로 표현가능하다. 즉, 주어진 데이터 y에 대해서 최적의 해 x를 찾는 것이 문제이며 이를 임의의 함수 $f(\cdot)$으로 표현가능하다. 데이터가 관찰되어 주어졌다는 표현을 "|"라는 기호를 이용함으로서 우리는 아래와 같은 함수에서 최적의 해를 찾는 문제를 정의할 수 있다. 위에서 언급한 다섯가지 문제들은 아래와 같은 형태로 재해석 가능하다.

1) $f($새로운 카드 사용의 정상여부|1년간의 카드사용 정보 및 새로운 카드사용 내역)
2) $f($적법한 사용자|인증요청이 들어온 지문과 인가된 사용자들의 지문정보들)

여기서 1)번 사례의 경우 "새로운 카드 사용의 정상여부"를 나타내는 미지수 값인 x는 $x\in\{$정상,비정상$\}$ 또는 $x\in\{0,\ 1\}$로 표현가능하다. 역시 관찰된 데이터에 대한 값 y는 "1년간의 카드 사용 정보 및 새로운 카드 사용 내용"을 나타낸다. 2번 사례의 경우 미지수값인 x는 "적법한 사용자"를 의미하므로 이에 대해서도 $x\in 0,\ 1$로 표현 가능하다. 이러한 규격화를 통해 바로 우리가 해결하고자 하는 목적함수를 얻을 수 있다.

$$\hat{x} = \arg\max f(x|y) \quad .$$

즉, 1번 사례의 경우, 새로운 카드 사용패턴이 기존의 정상 사용 패턴과 많이 달라서 $f(x=1|y)$의 값이 $f(x=0|y)$의 값보다 큰 경우에는 카드사기의 가능성이 높다고 볼 수 있으며, 이러한 경우 $x=1$의 값이 원하는 해로서 "비정상"이라는 결과 값을 주어야 한다. 그런데, 만약 $f(x|y)$값이 $f(x|y)\geq 0$ 이면서 $\int f(x|y)dx=1$을 만족하도록 만든다면 해당 함수는 사후확률분포(Posterior distribution)라는 형태로 말할 수 있으며 우리의 목적함수는 아래와 같이 작성이 가능하다.

$$\hat{x} = \arg\max p(x|y) \quad .$$

2.2. 베이지안 추론

이러한 목적함수를 해결하는 방법론으로 데이터 분석에 있어서 사용할 수 있는 기술들로 인공지능, 데이터 마이닝, 통계학, 통계신호처리 등이 대표적이며 기계학습과 데이터 마이닝 기술은 인공지능의 한분파로 간주가능하다. 그런데 이러한 데이터 분석들은 동일한 문제를 조금씩 다른 관점으로 풀고 있다. 하지만, 최근에 이러한 기술들의 공통된 원천기술로 여겨질수 있는 교집합으로서의 기술이 각광을 받아왔다. 그 기술이 바로 베이지안 추론 (Bayesian Inference)기술이다. 응용통계 및 통계신호처리 분야에서 주로 연구되어 왔던 Sequential Monte Carlo (SMC)나 Markov Chain Monte Carlo (MCMC)와 같은 몬테카를로 기법들(Monte Carlo)기법들은 대다수가 베이지안 추론기술에 기반하며 기계학습에서 주로 사용되어왔던 사후확률(Posterior)을 극대화시키는 최고치 값을 함수적 근사값으로 찾는 Variational Bayes 기반의 MAP 추정 기법이 그것들이며 최근에는 베이지안 재귀적 신경망(Bayesian Recurrent Neural Network [8])와 각종 베이지안 딥러닝 (Bayesian Deep Learning [3]) 기술들이 등장하고 있다. 이러한 다양한 베이지안 기술들은 현재 다양한 영역에서 기존의 추론 기술인 Frequentists' inference에 비해서 유용한 장점을 보였기에 그 사용이 증가하고 있다.

이러한 학문적 유용성을 지닌 베이지안 추론은 아래의 베이즈 규칙(Bayes Rule)에 기반한다.

$$p(x|y) = \frac{p(y,x)}{p(y)} = \frac{p(y|x)p(x)}{p(y)}$$

이러한 값들은 아래의 명칭으로 불리어진다.
-$p(x|y)$: 사후확률분포 (Posterior)
-$p(x,y)$: 결합 확률 분포(Joint distribution)
-$p(y|x)$: 우도함수(Likelihood)

-$p(y)$: 증거(Evidence/Marginal Likelihood)
-$p(x)$: 선지식, 사전확률(Prior distribution)

위식에서 볼수 있듯이, 사후확률은 우도함수와 사전확률값의 곱으로 표현되며 증거값에 의해서 정규화를 거친다는 것을 알 수 있다.

2.3. 베이지안 추론에 의한 문제 풀이

다시 2.1장에서 논의를 했던 문제 풀이로 돌아가 보자. 2.1장에서는 $f(x|y)$와 $p(x|y)$를 이미 언급하였다. 하지만, 실질적으로 원하는 해를 얻는 방식에는 아래와 같은 두 가지 형태중 하나로 표현가능하다.

- 우도함수(Likelihood): $p(y|x)$
- 사후확률분포(Posterior): $p(x|y)$

위의 두 가지 함수는 상당히 재미있는 특징을 보인다. 우도함수(Likelihood)는 얼마나 모델이 데이터에 적합하냐를 측정하는 함수라고 볼 수 있다. 즉, 임의의 해 x가 주어질 때 그해로부터 관찰 값인 y가 만들어질 확률 값을 나타낸다. 이와 달리, $p(x|y)$는 사후확률 값을 나타내는 것으로 관찰 값 y가 주어졌을 때 x가 나올 확률 값을 나타낸다. 우리가 주목해야 하는 것은 바로 이 사후확률 값이다. 그래서 2.1장에서 우도함수보다도 이 사후확률 값을 먼저 언급하였다. 이 사후확률 값은 아래와 같은 베이즈 규칙에 의해서 변형가능하다.

$$p(x|y) = \frac{p(y|x)p(x)}{p(y)} \propto p(y|x)p(x)$$

이식에서 알 수 있듯이, $p(x|y)$는 $p(y|x)$, $p(x)$, 그리고 $p(y)$로 구성되며 여기서 $p(y)$는 어떠한 불확실성도 없으므로 $p(x|y)$는 오직 $p(y|x)$와 $p(x)$에 의존함을 알 수 있다. 즉, 사후확률 값은 우도함수 값에 추가적으로 $p(x)$가 존재하는데 바로 이것이 베이지안 추론의 핵심인 사전확률 값인 $p(x)$이며 이 사전확률 값은 데이터나 관찰 값과는 어떠한 관련도 없는 함수 값이다.

하지만, 문제를 풀다보면 예상치 못한 변수들에 의해서 어려움을 겪게 된다. 즉, 주어진 관찰 값 y와 해결하고자 하는 해결안 x가 직접적으로 연결되어 있지 않은 경우 직접적인 해결이 쉽지 않다. 이러한 경우 아래와 같은 형태의 방식으로 사후 확률 값을 얻어야 한다.

$$p(x|y) = \int p(x,\theta|y)d\theta = \int p(x|y,\theta)p(\theta|y)d\theta$$

이 식은 x와 y가 θ라는 변수에 의해서 연결되어 있다고 보며 이러한 θ를 모든 경우에서 고려함으로서 우리의 목적함수인 사후확률(posterior)을 구할 수 있다고 말한다.

결론적으로 베이지안 추론은 $p(x|y)$을 찾아내는 과정이다. 불확실성을 갖는 미지수 (random variable)인 x에 대해서 [그림 1]과 같은 일종의 확률밀도함수 (probability density function, pdf)을 찾아내는 것이 바로 베이지안 추론의 목적이라고 볼 수 있다. 왜냐하면, 이러한 분포를 알아냄과 동시에 해결하고자 하는 모든 종류의 문제를 손쉽게 해결가능하기 때문이다.

가장 일례로, 사후확률분포를 안다면 아래와 같은 일반적인 통계치들을 용이하게 추출가능하다.

- 평균: $E(x|y) = \int xp(x|y)dx = \mu$
- 분산: $V(x|y) = \int (x-\mu)^2 p(x|y)dx$

또한 기계학습의 많은 문제들이 주로 이용하는 최적의 해를 찾는 문제는 아래와 같이 사후확률값 (posterior) 의 최대값을 갖게 하는 해를 찾는 문제와 동일하다. 2.1장에서 이미 언급한 이러한 추정을 MAP (Maximum A Posterior) 추정이라고 한다.

- 최대값: $\hat{x} = \arg\max p(x|y)$

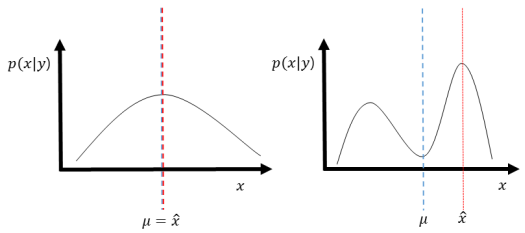
〔그림 1〕 Marginal posterior, $p(x|y)$

단, 인공지능에서 가능한 최적의 해를 찾기 위해서 위에서 언급한 평균값을 이용한 해와 최대값을 이용한 추정을 혼용해서 사용하는데 이때 조심해야 하는 경우가 존재한다. 아래 [그림 2]와 같이 상호 대칭인 구조를 갖는 사후확률의 경우에는 평균추정과 최대 추정간에 유사한 값을 보이는 반면 비대칭형이거나 불규칙적인 구조를 갖는 사후확률분포에서는 평균추정과 최대추정이 전혀 다른 값을 나타냄을 알 수 있다.

이러한 유용성은 더욱이 우리가 원하는 해x의 임의의 변환이 가해진 경우에도 이루어진다. 즉, 아래와 같은 수식을 얻을 수 있다.

$$E(h(x)|y) = \int h(x)p(x|y)\,dx$$

즉, 베이지안 추론에서는 확률 분포를 찾아냄으로서 다양한 연산이 가능해짐을 알 수 있다.



〔그림 2〕 평균추정(Mean estimate)과 최대값추정(MAP estimate)의 차이

2.4. 기계학습에서의 학습과 추론 개념 소개

2.1장에서 설명한 사후확률분포(Posterior distribution)을 구할 때 숨겨져 있는 모델변수들인 θ를 고려해야 하는 것을 살펴보았다. 다시 그 관계식을 작성하면 아래와 같다.

$$p(x|y) = \int p(x,\theta|y)d\theta = \int p(x|y,\theta)p(\theta|y)d\theta$$

이식을 자세히 살펴보면 결합 사후 확률 값(Joint posterior distribution)인 $p(x,\theta|y)$에서 모든 θ공간에서 적분을 하는 문제이다. 그런데, 이러한 적분이 x,θ,y간의 관계가 비선형적으로 복잡한 경우에는 실질적으로 $p(x|y)$를 깔끔한 수식으로 표현하기가 쉽지 않다. 이러한 이유로 $p(x|y)$를 구하기 위해서 응용통계학과 통계신호처리 분야에서는 점추정 기법(Point Estimate)인 몬테카를로 (Monte Carlo)기법을 이용하여 적분을 하나의 Dirac function 형태로 쪼개서 문제를 해결한다. 즉, 모든 θ 공간에서 상당히 큰 값인 N개의 샘플을 뽑아낸다고 가정하자. 그리고 이러한 샘플들이 $p(\theta|y)$로부터 나온다고 말한다면 아래와 같은 수식을 얻을 수 있다.

$$\theta^{(i)} \sim p(\theta|y), i = 1,2,\cdots,N$$

즉, 위와 같은 방식은 샘플링기법은 해당 분포를 아래와 같은 방식으로 근사값으로 표현가능하다.

$$p(\theta|y) \approx \frac{1}{N}\sum_{i=1}^{N}\delta_{\theta^{(i)}}(\theta)\,.$$

이러한 점추정기법에 근거하여 우리는 원래의 추정 분포 역시 찾아낼 수 있다.

$$\begin{aligned} p(x|y) &= \int p(x|y,\theta)p(\theta|y)d\theta \\ &\approx \int p(x|y,\theta)\frac{1}{N}\sum_{i=1}^{N}\delta_{\theta^{(i)}}(\theta)d\theta \\ &= \frac{1}{N}\sum_{i=1}^{N}p(x|y,\theta^{(i)}) \end{aligned}$$

하지만, 이러한 몬테카를로 (Monte Carlo) 방식은 비록 모든 공간에서의 θ를 만들어내지는 않는다는 장점은 있으나 여전히 충분히 큰 N을 사용해야 한다는 어려움이 존재하기에 점 추정 기법이 아닌 함수적 근사법(Functional approximation) 기반 알고리즘을 접근하게 되었다. 즉, 적분에서 사용되는 $p(\theta|y)$에 대해서 가장 유사한 하지만 다루기 쉬우면서 수학적으로 닫힌 형식을 갖는 (Numerically closed form) 함수로 근사화 한뒤에 아래와 같이 해결할 수 있다.

$$\begin{aligned} p(x|y) &= \int p(x|y,\theta)p(\theta|y)d\theta \\ &\approx \int p(x|y,\theta)q(\theta)d\theta \\ &= \tilde{p}(x|y) \end{aligned}$$

하지만, 이러한 방식 역시 근사화하는 함수 $q(\theta) \approx p(\theta|y)$ 로서 근사화를 통해서 유실되는 정보가 적어야 한다는 단점이 있으며 이러한 제대로 된 함수를 찾는 것이 상당히 어려운 문제로 여겨지고 있고 기계학습 분야에서 Variational Inference, Gaussian Approximation, Expectation Propagation Approximation 과 같은 다양한 방식들에 의해서 만들어진 Functional Approximation 기법들이 제안되었으나 데이터와 문제마다 그 성능의 변화가 크기에 따라 민감하게 사용되어야 한다는 단점은 유효하다고 볼 수 있다.

이러한 이유로 기계학습 분야에서는 가장 간단한 방식중의 하나로는 적분 방식을 포기하는 것이다. 즉, 모든 θ공간에서의 적분 값을 이용하기보다는 가장 유효한 결과 값을 제시하는 특정 θ값에 대해서만 고려함으로서 적분에서의 어려움을 제거하는 방식이다. 즉, 특정 $\hat{\theta}$를 얻은 후에 이 특정 $\hat{\theta}$를 데이터 y처럼 관찰되었다고 가정하고 문제를 푸는 방식으로 수식으로 표현하면 아래와 같다.

1) $\hat{\theta} = \arg\max p(\theta|y)$
2) $\hat{x} = \arg\max p(x|y) \approx \arg\max p(x|y,\hat{\theta})$

이러한 Two step방법은 직접 적분으로 최적의 해를 구함이 어렵기에 발생한 근사방법인데 주목할 점은 바로 이 방식이 거의 모든 기계학습 분야에서 사용하고 있는 학습(테스트)과 추론 기법이라는 것이다. 즉, 1)번 과정에서 해를 구하기 위해서 필요한 다양한 모델변수들(θ)을 주어진 데이터로 학습을 한 뒤에 이렇게 학습과정으로 얻어진 $\hat{\theta}$를 이용하여 새로운 질의에 바로 즉각적으로 예측이나 분류 등을 실행할 수 있다.

Ⅲ. 정보보호 이슈로의 접목

3.1. 악성코드 분류 (Malware detection)

악성 코드 분석은 가장 대표적인 정보보호 이슈이다. 최근에 인공지능의 발달로 인해서 정보보호에서 AI기반의 악성코드 분석이 이슈화하고 있지만, 사실은 이미 정보보호분야에서 인공지능이 몇 십년 전부터 사용되어왔던 분야중 하나가 바로 악성코드분석이었다. 하지만, 최근에는 시그네쳐(Signature) 기반의 정적분석보다는 실행환경에서의 API call이나 동적패턴을 특징점들(Feature)을 선택하고 정제하여 고도의 기계학습 기술을 이용하여 분류성능을 높이고 있다는 점에서 기존의 수동적인 인공지능 기반 분류보다 최근에 이슈로 되고 있는 MAX와 같은 인공지능 기반 악성코드 분류 기술들은 여러 가지 면에서 유의미하다고 말할 수 있다.

이러한 악성코드분류기술은 기계학습기반으로 구분을 짓는다면 일종의 분류(Classification)기술이며 이러한 분류기술로 Support Vector Machine, 인공신경망(Artificial Neural Network), 딥러닝(Deep Learning), Logistic Classification, 선형분류(Linear Classification), KNN classifier등 다양한 기법들이 사용될 수 있다.

하지만, 여기서 중요한 것은 이러한 모든 기술들이 베이지안 추론 표형방법을 이용하면 하나의 식으로 아래와 같이 기술이 가능하다는 점이다.

$$\hat{x} = \arg\max p(x|y) \quad , \; x \in \{0,1\}$$

여기서 y는 주어진 악성코드와 정상 파일들의 데이터 셋들이다.

3.2. FDS as outlier detection

카드사기 범죄를 찾는 기술인 FDS (Fraud Detection System)은 현재 금융권에서 가장 중요하게 다루는 정보보호의 한 분야이다. 단순하게 생각하면 FDS역시 일종의 분류 (Classification)기술로 여겨질 수 있으나, 내용적으로는 꼭 그렇지가 않다. 정상상태의 데이터에 비해서 비정상상태의 데이터가 너무나 현저히 작기 때문이며 어떤 패턴으로 일어날지 알려지지 않은 상황 때문이다[5]. 그래서 FDS는 단순한 분류기술을 쓰기에 앞서 세 가지 방법이 고려되어야 한다.

1) 정상데이터와 아주 희박하게 발생하는 비정상 데이터들의 불균형적인 데이터 수에서 발생하는 문제 해결이 관건으로 이러한 불균형데이터(Imbalanced data)에 대한 이슈가 먼저 해결되어야 한다. 이러한 것을 해결하는 방법으로 우선 대량의 데이터를 갖고 있는 그룹

의 데이터를 적은 양의 데이터 그룹의 크기만큼 데이터 양을 줄이는 방법과 반대로 적은 그룹의 데이터들을 강제로 데이터를 가상으로 만들어내어 Balanced data로 만드는 방법이 있다.

2) 이러한 불균형데이터(Imbalanced data)들에 대해서 사전확률 (Prior distribution)을 이용하여 데이터의 불균형 특징을 약화시킬 수 있다. 이러한 데이터에서 발생하는 문제 역시 베이지안 추론 기술로 효과적으로 해결할 수 있다. 즉, p(x=1)과 p(x=0)에 대한 이론적인 근거를 사전확률 값으로 적용함으로서 불균형 문제를 해결할 수 있다.

3) 마지막으로 단순한 분류기법이 아닌 이상치탐지(Outlier Detection)문제로 해당 문제를 접근하는 방식이 필요하다[2]. 즉, 카드사기 사건과 같은 현상은 확률적으로 일어나기 어려운 현상이라는 점에서 극값이론(Extreme Value Theorem, EVT)과 같은 방법을 이용함으로서 해결가능하다. 흥미로운 점은 이 EVT개념 역시 베이지안 기술로 효과적으로 기술가능하다 [1].

이러하듯, 세가지 방안을 이용하여 FDS기술을 만들 수 있는데 이 때 사용되는 기술들 모두가 베이지안 기법으로 효과적으로 만들어질 수 있다는 점이 중요하다 [6,7].

Ⅳ. 결 론

본 논문에서는 최근의 인공지능과 기계학습에서 주축중 하나를 이루는 베이지안 기술을 소개하는데 그 목적을 두었다. 특히 기계학습에서 데이터 분석시에 데이터를 이용하여 모델변수들을 학습한 후에 새로운 질의에 대한 결과 값을 테스트 과정을 거쳐서 얻어내느지에 대한 방법론을 제시하였다. 또한 이후에 정보보호분야에서 대표적으로 사용되는 악성코드 분류 및 카드사기 탐지시스템 (FDS)에 대한 베이지안 관점에서의 해석도 소개하였다.

참 고 문 헌

[1] S. J. Roberts, “Novelty Detection using Extreme Value Statistics,” IEE Proceedings – Vision, Image, and Signal Processing, Vol. 146, No. 3, pp. 124-129, June 1999.

[2] K. Chaloner and R. Brant, “A Bayesian approach to outlier detection and residual analysis,” Biometrika, Vol. 75, No. 4, pp. 651-659, 1988.

[3] Bayesian Deep Learning, http://bayesiandeeplearning.org/, NIPS 2016 Workshop

[4] MAX, Saint Security, 2017, http://www.stsc.com/max.html

[5] R. J. Bolton and D. J. Hand, “Statistical Fraud Detection: A Review,” Statistical Science, Vol. 17, No. 3, pp. 235-249, 2002.

[6] S. Maes, K. Tuyls, B. Vanschoenwinkel, and B. Manderick, “Credit Card Fraud Detection Using Bayesian and Neural Networks”, Proceedings of the 1st International Naiso congress on Neuro Fuzzy Technologies, pp. 261－270, 1993

[7] M. S. Shotwell and E. H. Slate, “Bayesian Outlier Detection with Dirichlet Process Mixtures”, Bayesian Analysis, Vol 6, No. 4, pp. 665-690, 2011

[8] M. Fortunato, C. Blundell, and O. Vinyals, “Bayesian Recurrent Neural Networks”, arXiv, Vol abs/1704.02798, , 2017

〈저자소개〉

윤 지 원 (Yoon, Ji Won)
정회원

2008년 11월 : University of Cambridge, 전자공학과 박사졸업
2008년 2월~2009년 5월 : University of Oxford, 로봇연구소 박사후과정
2009년 5월～2011년 8월 : University of Dublin, 통계학과 연구원 및 강사
2011년 7월～2012년 8월 : IBM연구소 연구원
2012년 9월～현재 : 고려대학교 사이버국방학과, 정보보호대학원 부교수
관심분야 : 신호정보처리, 응용통계, 도감청 탐지 기술, Open Source Intelligence