방송기술의 클라우드 전략
Part 6. 인공지능과 클라우드

글.
김기완 AWS 솔루션즈 아키텍트

연재목록

혁신은 새로운 기술로부터 시작되기도 하지만, 사람들의 상상력으로부터 시작되기도 한다. 사람과 자유롭게 대화하며 필요한 일들을 척척 해내는 기계에 대한 호기심은 인공지능 기술의 발전을 가져왔다. 아마존은 기계학습을 통하여 대화로 소통하는 '아마존 에코'를 통하여 음성 기반 UI로의 발전을 선도하고 있으며, 인공지능 드론을 통한 배송, 컴퓨터 비전과 인공지능을 접목하여 무인 편의점인 '아마존 고'까지 다양한 인공지능 서비스들을 출시하고 있다.

국내에서도 음성기반 개인비서 서비스에 대한 관심이 높아지고 있고, 실제로 다양한 제품들이 출시되고 있다. 제조 분야에서는 제조된 제품에 대한 이미지 분석을 통해 불량 여부 판단이나 여러 센서로부터 모인 데이터에 대한 분석을 통해, 불량률을 예측하는 데 활용되고 있고, 금융 분야에서는 비정상적인 금융 거래 분석에 인공지능을 사용하고 있으며, 많은 전자상거래 기업들도 고객 맞춤형 추천 서비스 및 구매 가능성 분석을 위해 인공지능을 도입하려 하고 있다.

방송 및 미디어 산업도 예외는 아니다. 넷플릭스는 75%의 콘텐츠가 추천을 통하여 소비되고 있다고 밝혔다.[1] 이를 위해 넷플릭스는 인공지능 및 기계학습을 활용하고 있다.[2] 구글의 YouTube도 콘텐츠 추천을 위하여 신경망 기반의 머신 러닝 기법을 사용하고 있다.[3]

일반적으로 인공지능을 효과적으로 사용하기 위해서는 딥 러닝을 비롯한 다양한 기계학습 기법들에 대한 이해 및 경험, 그리고 숙련된 전문가가 필요하다. 하지만 클라우드는 기계학습을 비용 효율적이고 빠르게 수행하는 데 도움을 준다. 클라우드는 현실적으로 무제한 제공되는 컴퓨팅 자원을 활용해 빠른 학습이 가능하도록 하고, 효과적인 학습을 위한 모델을 완전 관리형으로 제공한다. 이를 통해 데이터만으로 빠르게 인공지능을 사용할 수 있다. 또한 인공지능 및 기계학습에 대한 전문지식이 없는 개발자들도 서비스 형태로 사용할 수 있는 인공지능 서비스(이미지 분석, 음성 인식, 챗봇 서비스)들이 출시되고 있다.

이번 호에서는 인공지능의 기술에 대하여 간단히 살펴보고, 클라우드에서 인공지능을 시작할 수 있는 방법에 대해서 알아보자.

---

1. https://medium.com/netflix-techblog/netflix-recommendations-beyond-the-5-stars-part-1-55838468f429
2. https://medium.com/netflix-techblog/learning-a-personalized-homepage-aa8ec670359a
3. https://static.googleusercontent.com/media/research.google.com/en//pubs/archive/45530.pdf

김기완
AWS 솔루션즈 아키텍트

인공지능(Artificial Intelligence), 기계학습(Machine Learning) 이란?

데이터로부터 의미를 얻어 내는 작업은 낯선 일은 아니다. OTT 기반의 미디어 서비스에서 개인의 특징을 고려한 미디어 추천 서비스 구성을 예로 들어 보자. 쉽게 생각할 수 있는 것은 사람들을 몇 개의 그룹으로 나누는 작업이다. 이러한 그룹별 분류를 하기 위해, 일단 사용자들의 특징(attribute)들을 나열한 하나의 데이터베이스 테이블을 만들고, 사용자들의 모든 특징을 저장한 후 SQL 쿼리 등을 통하여 각 사용자들을 특정 군으로 분류한다. 이를 위해 주로 비즈니스 인텔리전스(Business Intelligence, BI) 도구들을 통하여 작업을 하게 된다.

이러한 작업은 데이터 과학자(Data Scientist)라 부르는 사람들이 경험 및 데이터 분석을 통해 수행하여 왔다. 즉, 20대 여성을 하나의 군으로 분류하고, 30대 전문직 남성을 하나의 군으로 분류하는 등 사용자의 특징들을 기반으로 군을 분류하게 된다. 이러한 작업은 사람의 통찰(insight)을 통해 이루어지게 되며, 과거 데이터에 의존하게 된다. 사용자들의 행동(behavior)이 변화하게 되면 새로운 통찰을 발견해 내야 하고, 대개 이러한 작업은 많은 시간이 걸리게 된다. 경우에 따라서는 과거에 없는 새로운 데이터를 수집해야 할 수도 있고, 이러한 경우 전체 서비스에 변경이 필요할 수 있다.

이를 단순한 수식으로 표현한다면 다음과 같다.

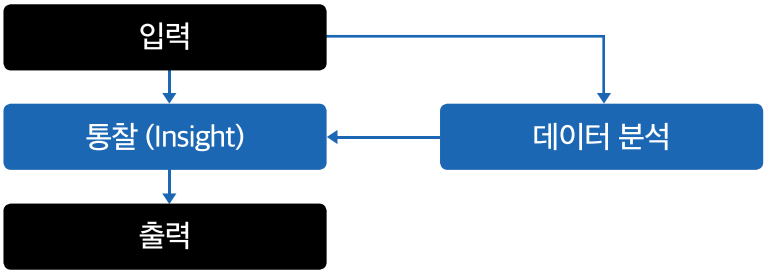

그림 1. 추천 과정

즉, 입력되는 데이터는 분석을 통해 '통찰'이 이루어지게 된다. 일단 일정량의 데이터 분석을 통해 통찰을 얻게 되면, 이러한 통찰이 비즈니스 로직에 반영된다. 이후에는 유저에 대한 정보가 입력되면 얻어진 통찰(대개 SQL 쿼리나 이에 준하는 업무 프로세스로 개발된다)을 기반으로, 유저를 특정한 군으로 분류하는 작업을 할 수 있다. 이러한 군에 따라 추천 콘텐츠를 제공한다면 원하는 시스템이 완성된다.

기계학습은 통찰을 얻는 과정을 사람의 힘이 아닌 기계를 이용하는 방식의 학습이다. 즉, 다음과 같은 과정을 거치게 된다.

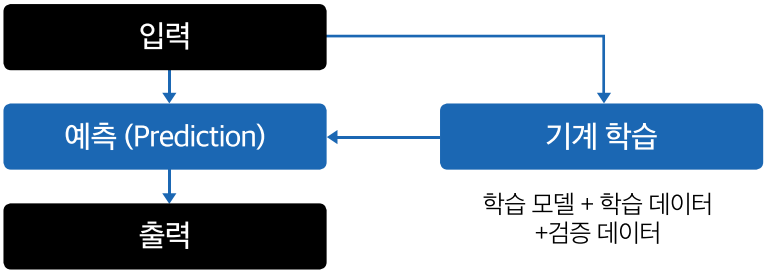

그림 2. 기계학습의 개요

입력되는 데이터 중 일부는 학습 데이터 및 검증 데이터로 나뉘게 되고, 구성되어 있는 학습 모델을 통해 예측 모델이 만들어진다. 검증 데이터를 통해 예측 모델의 정확성이 테스트되고, 설정한 임계값보다 높은 정확성을 갖는 예측 모델이 만들어지면, 이러한 예측 모델을 통해 새롭게 유입되는 입력값에 대한 출력이 만들어진다(사용자에 대한 분류 작업 등).
정리해 보면, 유입되는 데이터에 대하여 원하는 예측값을 출력하는 함수, 이 함수를 만들어 내는 과정을 기계학습이라 할 수 있다.

지도 학습 (Supervised Learning), 비지도 학습 (Unsupervised Learning)

이러한 기계학습은 지도 및 비지도 학습으로 나누어 볼 수 있다.

- 지도학습 : 학습 데이터는 입력과 기대되는 출력값의 쌍(벡터)으로 구성되어 있음. 즉, 학습 데이터의 개별 벡터는 정답을 가지고 있고, 이를 통해 정확성을 검증할 수 있음.
- 비지도학습 : 학습 데이터는 입력에 대한 올바른 출력값을 가지고 있지 않음. 여러 방법을 이용하여 출력 계층을 구성하여야 함.

지도학습에서는 학습에 사용되는 데이터에 입력값뿐 아니라 출력값이 포함되어 있다. 예를 들어 이미지를 통하여 해당 이미지에 존재하는 사물을 인식하는 기계학습을 진행하는 경우, 각 이미지가 해당 이미지에 대한 이름, 즉 레이블(label)을 가지고 있는 경우를 생각해 볼 수 있다. 즉, 고양이에 대한 이미지는 '고양이'라는 레이블을 가지고 있고, 이미지가 입력되었을 때 '고양이'라는 출력값이 나와야 한다.
비지도학습에는 이러한 출력 레이블이 존재하지 않는다. 예를 들어 미디어 추천 서비스를 제작할 때 일련의 사용자들이 비슷한 군에 존재한다는 것은 알 수 있으나, 이들에 대한 고정된 출력값이 없는 경우를 생각해 볼 수 있다.

학습 모델

기계학습이 주어진 입력값으로부터 출력값을 만들어내는 과정이라고 보면, 역시 가장 중요한 것은 학습 모델과 학습 데이터이다. 많은 기계학습 전문가들이 모든 문제를 해결할 수 있는 하나의 훌륭한 학습 모델을 개발하기 위하여 노력하고 있다. 하지만, 실제로는 문제에 따라 정확도가 높은 모델들이 다른 것이 일반적이다. 많이 사용되는 학습 모델에는 결정 트리 학습법(Decision Tree Learning), 신경망을 사용한 딥러닝(Deep Learning), 확률론에 근거한 베이지안 학습법(Bayesian Learning) 등이 있다.

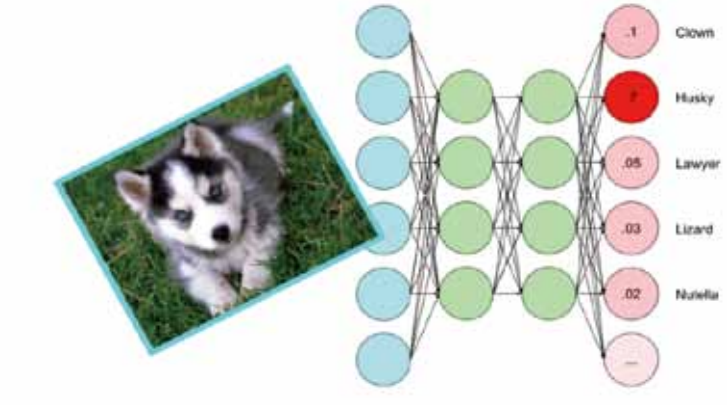

그림 3. 딥러닝에서 사용되는 신경망

최근 이미지 분석 및 음성 인식에 많이 사용되고 있는 딥러닝에 대해서 조금 더 자세히 살펴보자. 학습 데이터가 정형화되어 있고, 각각의 특징(feature)들이 데이터베이스 스키마 안에 하나의 필드(field)로 정의된 경우와는 달리, 이미지/음성 등의 비정형 데이

김기완
AWS 솔루션즈 아키텍트

터들은 그 안에서 특징을 찾아내는 것이 쉽지 않다. 딥러닝은 이미지나 음성 등의 데이터에서도 특징들을 찾아낼 수 있는 방법을 제공한다. 대표적인 딥러닝 기법의 하나인 신경망 네트워크는 다음과 같이 구성된다.
[그림 3]에서 보이는 것과 같이 신경망에 기초한 딥러닝은 입력 계층(하늘색 원)과 은닉 계층(Hidden Layer, 초록색 원), 그리고 출력 계층(붉은색 원)으로 구성된다. 예컨대 사진과 같은 입력 데이터가 입력 계층으로 들어오면, 데이터에 대한 맵이 만들어지고 은닉 계층에서 해당 이미지의 특징들이 학습을 통해서 생성된다.

특징이 은닉 계층에서 학습을 통해서 생성된다는 개념이 다소 어려울 수 있겠다. 쉬운 이해를 위해 딥러닝에서 사용되는 신경망은 실제 사람의 신경망을 본뜬 것이므로, 다음과 같이 생각해 볼 수 있을 것이다. 유명인의 사진을 보고, 그 사람의 이름이 무엇인지를 떠올리는 과정을 생각해 보자. 눈을 통해서 시각적인 자극이 들어오면(입력 계층) 뇌 속의 특정 뉴런들은 전기적 신호를 다양한 강도로 전파(은닉 계층)하게 되어 결국 뇌 속의 특정 기억 장치를 자극하게 된다(출력 계층). 그러면 머릿속에서는 해당 유명인이 누구인지를 알 수 있게 된다.

위의 과정에서 어떤 뉴런들이 어떤 강도로 신호를 전파하는가는 발견되는 특징에 따라 다르다. 신경망 중심의 딥러닝 기법에서는 하나하나의 은닉 노드를 다음과 같이 실제 뉴런과 비슷한 모양으로 형상화한다.

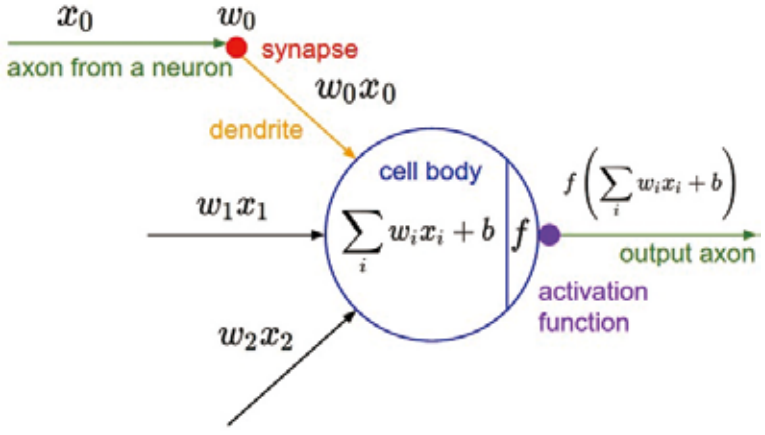

그림 4. 딥러닝에서 사용되는 단위 신경 세포

[그림 4]에서 해당 단위 신경 세포로 전달되는 입력값은 Xn이다. 이 신경 세포와 연결되어 있는 이전 단계의 단위 신경 세포의 개수만큼 입력값이 들어오게 되고, 각각의 입력값은 가중치 값을 가지고 있다(Wn). 그러면 단위세포는 각 입력값과 가중치의 곱의 합에 보정값을 더하여 활성화 함수(activation function)를 거친 출력값을 다음 노드로 전파하게 된다. 이러한 과정에서 특징(feature)은 가중치 값의 행렬(또는 벡터)로 구체화된다.

많은 학습 데이터들을 통해 은닉 계층에 있는 각각의 뉴런 세포들의 가중치 값들의 벡터가 만들어지면, 하나의 사용 가능한 딥러닝 신경망이 완성되고, 이제 새로운 입력에 대해서 원하는 출력 값을 생성할 수 있는 모델을 사용할 수 있게 된다.

클라우드에서의 기계학습

앞에서 간단히 기계학습의 개념과 신경망 구조의 딥러닝에 대해서 살펴보았다. 그러면, 이러한 기계학습을 클라우드에서 어떻게 사용할 수 있는지 AWS의 예를 들어 살펴보자. AWS의 경우 크게 세 가지 방법으로 기계학습을 활용할 수 있도록 돕고 있다.

딥러닝 전문가를 위한 AWS 인프라 사용

딥러닝을 많이 사용하고 있는 전문가들은 이미 Tensorflow, Caffe, Torch, MXNet 등의 딥러닝 프레임워크에 익숙하다. 이러한 딥러닝 전문가들을 위해서 AWS는 빠르게 딥러닝 프레임워크를 사용할 수 있도록 인프라 구성 템플릿(CloudFormation Template)을 제공하고 있다.[4]

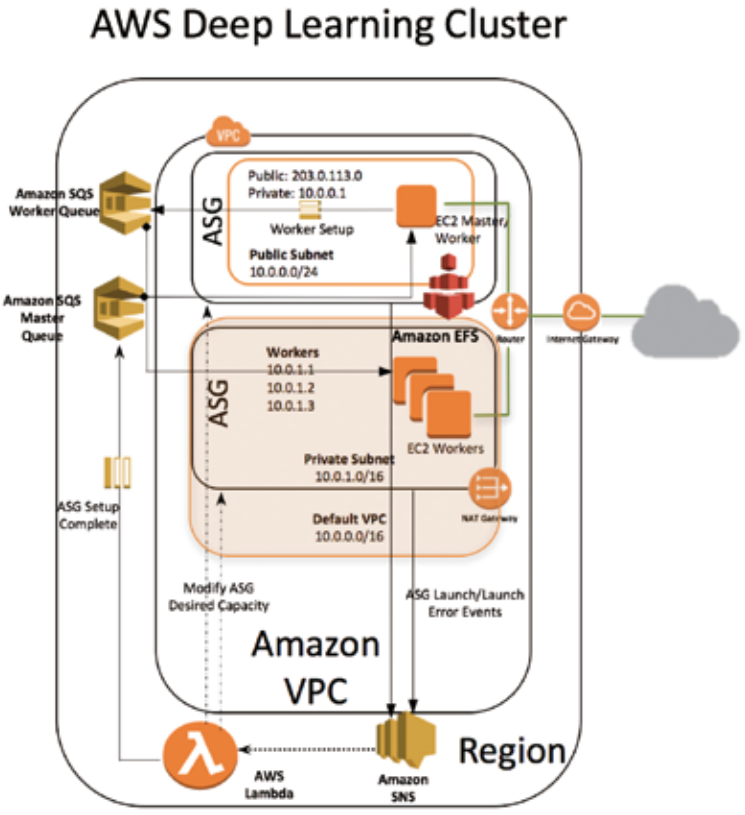

그림 5. AWS의 딥러닝 CloudFormation 템플릿

AWS의 딥러닝 CloudFormation 템플릿을 사용하면 GPU를 포함하여 원하는 자원들에 대한 설정만으로 빠르게 MXNet, 텐서플로우, Caffe, Theano, Torch 및 CNTK 등과 같은 딥러닝 프레임워크를 분산 환경에서 사용할 수 있도록, 인프라가 10여 분만에 배포될 수 있다. 반면 고가의 GPU를 장착한 물리적인 서버 환경을 구성하기 위해서는 많은 절차와 시간이 들게 될 뿐 아니라, 원하는 만큼 자원을 빠르게 사용하기도 어렵다.

물론 이러한 딥러닝 템플릿 이외에도 완전 관리형 하둡 서비스인 Amazon EMR(Elastic Map Reduce)을 통하여 Spark Machine Learning 혹은 Mahout 기반의 머신 러닝을 사용할 수 있다.

데이터 과학자를 위한 완전 관리형 기계학습 서비스 (Amazon Machine Learning)

AWS는 Amazon Machine Learning(AML) 서비스를 통해 기계학습 및 예측 서비스를 완전 관리형으로 제공한다. AML을 통해서 고객은 과거의 데이터를 학습용/검증용으로 나누어 AWS의 관리형 모델에 적재하고 원하는 정확도로 학습을 수행하여 이를 실제 업무 환경에 적용할 수 있다.

4. https://medium.com/netflix-techblog/netflix-recommendations-beyond-the-5-stars-part-1-55838468f429

김기완
AWS 솔루션즈 아키텍트

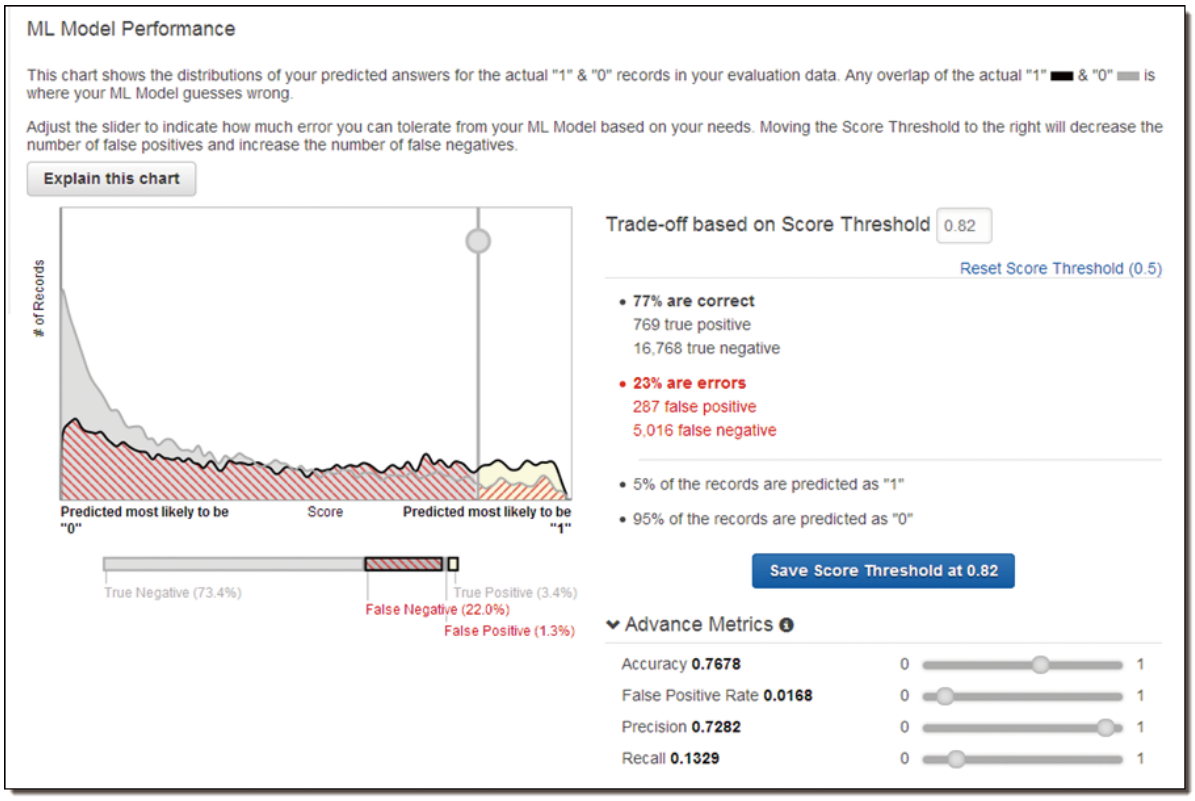

그림 6. Amazon Machine Learning 서비스

Amazon AI 서비스

AWS는 이미지 분석 서비스인 Amazon Rokognition, 텍스트를 읽어주는 TTS(Text To Speech) 서비스인 Amazon Polly, 그리고 챗봇 서비스를 빠르게 개발할 수 있도록 도와주는 Amazon Lex 서비스를 제공하고 있다.

그림 7. Amazon Rekognition

[그림 7]과 같이 Amazon Rekognition은 이미지 분석을 통해 이미지에 존재하는 레이블들을 확률과 함께 얻을 수 있도록 한다. 이미지를 AWS API/SDK를 통하여 Rekognition 서비스에 전달하면, 이미지에 존재하는 각 레이블들이 확률과 함께 JSON 포맷으로 전달된다. 이러한 값들은 애플리케이션에서 다양하게 활용될 수 있다. 또한 Amazon Rekognition 서비스는 얼굴 분석에도 사용될 수 있다.

그림 8. Amazon Rekognition의 얼굴 분석 서비스

Amazon Rekognition은 이미지에 존재하는 얼굴을 특성값과 함께 저장할 수 있고, 이렇게 분류된 얼굴들을 새로 입력되는 얼굴과 비교하여 동일 얼굴을 감지할 수 있다. 이러한 기술은 다양하게 응용될 수 있다. 예를 들어 동영상에서 이미지를 추출하고 추출된 이미지를 분석하여 동영상 안에서 특정인이 나오는 장면들을 찾아낼 수 있다.

연재를 마치며

지난 1월부터 6월까지, 총 6개월에 걸쳐 AWS를 중심으로 미디어 산업에서의 클라우드 활용에 대해서 살펴보았다. 클라우드는 단순한 웹 서비스부터 기계학습에 이르기까지 다양한 IT 분야에서 활용될 수 있다. 또한 미디어 트랜스코딩, 렌더링, 라이브 스트리밍에 이르기까지, 디지털 미디어 서비스의 핵심 기능들은 클라우드상에서 빠르게 구축할 수 있게 되었다. 지난 4월 26일에 진행된 NAB Show의 슈퍼세션에서는, 우주 상공에서의 4K 라이브 스트리밍이 AWS 서비스를 통해 전 세계에 서비스되기도 했다. 이제 클라우드는 더 이상 가능성의 영역에 머무르고 있는 기술 개념이 아니라, 실제 수많은 방송 및 미디어 업체들이 활용하고 있는 핵심 플랫폼으로 자리 잡았다.

클라우드 도입은 단순하게 하나의 솔루션이나 서비스를 도입하는 것과는 차원이 다르다. 클라우드에서는 데이터센터에서 이루어지는 모든 작업이 구현 가능하며, 환경의 변화에 따라 유연하게 비용 효율적인 미디어 서비스를 구축할 수 있다. 오늘날 소비자는 과거의 공급자 중심의 미디어 서비스에 더는 머물러 있으려 하지 않는다. 급변하는 소비자의 요구에 맞추어 빠르게 새로운 서비스를 제공하기 위해서는, 클라우드의 도입을 더 이상 늦추거나 미뤄서는 안 될 것이다. 또한 클라우드 기술을 잘 활용하기 위해서는 오랜 시간과 축적된 경험에 기반한 클라우드 역량이 필요하다. 이제 작은 프로젝트를 통해, 클라우드에 대한 지식과 경험을 직접 얻어보고, 이러한 작은 단계에서부터 클라우드로의 여정을 시작해 보자.

김기완 솔루션즈 아키텍트는 IBM에서 15년 간 여러 기술 분야 업무를 수행하였으며 메인프레임 및 유닉스 운영체제, 데이터베이스 및 각종 미들웨어와 다양한 애플리케이션에 이르기까지 엔터프라이즈 고객들과 다양한 환경에서 많은 경험을 가지고 있다. 폭넓은 기반 지식과 이러한 고객 경험을 바탕으로 방송/미디어 그리고 다양한 엔터프라이즈 고객들의 AWS로의 여행을 돕고 있다.