



## 『넷플릭스 기술 연구회』 2부

EBS 기술인협회 스터디  
『넷플릭스 기술 연구회』

지난 8월호에 처음 실렸던 1부에서는 EBS 기술인협회에서 진행하고 있는 스터디 지원 사업을 소개했고, ‘넷플릭스의 오픈소스 정책’, ‘머신러닝을 통한 스트리밍 품질 향상’ 두 가지 기사에 대해 포스팅했습니다. 이번 2부에서는 ‘차세대 이미지 코딩을 위한 AVIF’, ‘예측 모델링으로 넷플릭스의 콘텐츠 품질관리 최적화’ 두 가지에 대해서 포스팅하도록 하겠습니다. 해당 포스팅이 이 글을 읽는 독자들에게 도움이 되는 자료가 되었으면 합니다.

### 차세대 이미지 코딩을 위한 AVIF

#### 원본 정보

|      |                                                                                                                                                                           |                                                                                      |
|------|---------------------------------------------------------------------------------------------------------------------------------------------------------------------------|--------------------------------------------------------------------------------------|
| 제목   | AVIF for Next Generation Image coding<br>(차세대 이미지 코딩을 위한 AVIF)                                                                                                            |  |
| Link | <a href="https://netflixtechblog.com/avif-for-next-generation-image-coding-b1d75675fe4">https://netflixtechblog.com/avif-for-next-generation-image-coding-b1d75675fe4</a> |                                                                                      |

그림 1. 원문 QR-Code

이번 포스팅에서는 비디오와 이미지를 압축하는 방식과 품질에 대해 간략하게 말씀드리고자 합니다.

#### Image compression at Netflix

| 원문                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                          | 번역                                                                                                                                                                                                                                                                                                                                                                                                                         |
|-------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------|----------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------|
| <p>Netflix is enjoyed by its members on a variety of devices — smart TVs, phones, tablets, personal computers and streaming devices connected to TV screens.</p> <p>The user interface (UI), intended for browsing the catalog and serving up recommendations, is rich in images and graphics across all device categories.</p> <p>...</p> <p>The goal is to have the compressed image look as close to the original as possible while reducing the number of bytes required.</p> <p>Given the image-heavy nature of the UI, <u>compressing these images well is of primary importance.</u></p> <p>This involves picking, among other things, the right combination of color subsampling, codec, encoder parameters and encoding resolution.</p> <p>...</p> | <p>넷플릭스는 스마트 TV, 휴대폰, 태블릿, 개인용 컴퓨터 및 TV 스크린에 연결된 스트리밍 장치 등 다양한 기기에서 회원들에게 이용되고 있습니다.</p> <p>카탈로그를 탐색하고 추천을 제공하기 위해 고안된 사용자 인터페이스(UI)는 모든 디바이스에서 이미지와 그래픽이 풍부하게 사용됩니다.</p> <p>...</p> <p>압축 이미지의 목표는 가능한 한 원본과 가장 비슷하게 보이면서 필요한 바이트 수를 줄이는 것입니다.</p> <p>UI가 이미지 위주로 구성되어 있는 점을 고려하면, 이러한 <u>적절한 방법으로 압축된 이미지를 사용하는 것이 가장 중요합니다.</u></p> <p>이것에는 적절한 컬러 서브샘플링, 코덱, 인코더 매개변수 및 인코딩 해상도의 적절한 조합을 선택하는 것이 포함됩니다.</p> <p>...</p> |

본문에서 언급한 것처럼 다양한 압축 방식이 있는 만큼, 이미지와 비디오를 용도에 맞추어 적절히 사용하는 것이 중요합니다. 내용의 이해를 쉽게 하기 위해서 비디오 압축 방식에 대해 먼저 말씀드리겠습니다.

### Video Compression

비디오를 압축하는 방식으로는 크게 IntraFrame, InterFrame 두 가지 방식으로 나눌 수 있습니다. 이 두 방식의 차이는 구성하고 있는 프레임의 종류에서 나타납니다. 먼저 IntraFrame 방식은 모든 프레임들이 I-프레임으로 구성되어 있습니다. I-프레임은 프레임 자체가 모든 정보를 가지고 있는 하나의 이미지 파일이라고 생각할 수 있습니다.

다음으로 InterFrame 방식은 I(Intra), P(Predictive), B(Bidirectional)-프레임 총 3종류의 프레임으로 구성되어 있습니다. P-프레임은 순방향 예측(이전의 I-프레임 참조)을 통해 공통부분은 뒤 프레임에서 제거하고 변경된 부분만 저장하는 방식입니다. 그리고 B-프레임은 양방향 예측(이전과 이후의 I 또는 P프레임 참조)을 통해 역시 변경된 부분만 저장하는 방식입니다.

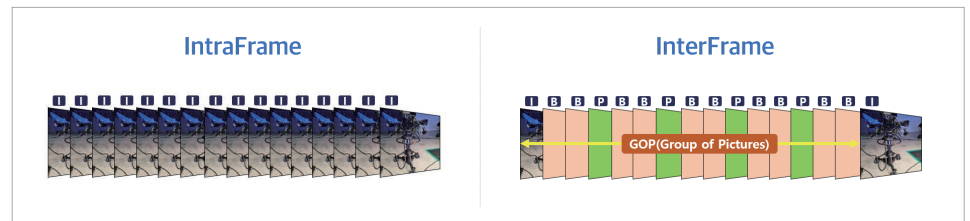


그림 2. Intraframe, Interframe 비교

이 3가지 종류의 프레임을 몇 개나 사용하느냐에 따라서 GOP(Group of Pictures)의 길이가 달라집니다. 일반적으로 TV 전송에서 우리가 사용하는 NTSC 신호는 15프레임을 하나의 그룹으로, PAL 신호는 12프레임을 하나의 그룹으로 사용합니다.

### Brief overview of image coding formats

기사 본문에서 언급한 JPG, Webp 등을 포함해 다른 형식의 특징도 간략하게 소개해 드리려고 합니다.

첫 번째 형식은 **JPG(Joint Photographic experts Group)**입니다. JPG는 손실 압축이며 알파 채널(이미지나 비디오의 각 픽셀의 투명도 저장) 및 애니메이션을 지원하지 않습니다.

두 번째는 **PNG(Portable Network Graphics)**입니다. PNG는 이미지의 정보는 삭제하지 않으면서 크기를 줄이는 무손실 압축으로 JPG보다 많은 색을 표현할 수 있으며 알파 채널을 지원하지 않지만, 애니메이션은 지원하지 않습니다. 우리는 현재 이미지를 저장할 때 이 두 가지 형식을 가장 많이 사용합니다.

세 번째는 **GIF(Graphics Interchange Format)**입니다. 이 형식 역시 무손실 압축이며 알파 채널을 지원하지 않습니다. 우리가 애니메이션을 저장할 때 현재 가장 많이 사용하는 형식입니다.

다음으로는 위의 3가지 형식을 대체하기 위해 만들어진 Webp, HEIF, AVIF에 대해 알아보겠습니다.

**Webp**는 웹사이트에서 이미지가 로딩되는 시간을 단축하기 위해 2010년 Google에서 출시한 형식입니다. 가장 큰 이점은 손실, 무손실 압축을 둘 다 지원하며 손실 압축을 사용하면 JPG보다 30%, 무손실 압축을 사용하면 PNG보다 2~30% 정도로 비슷한 품질을 적은 용량으로 사용할 수 있습니다.

**HEIC(High Efficiency Image Container)**는 2017년 애플에서 표준 포맷으로 채택하며 사용되었습니다. 용량이 JPG의 50% 정도이며 높은 품질의 이미지로 저장할 수 있습니다. 하지만 애플 외의 장비나 소프트웨어에서는 호환성이 떨어진다는 단점이 있습니다.

마지막은 본문에서 소개하는 **AVIF(AV1 Image File Format)**입니다. AVIF는 아마존, 애플, 넷플릭스 등이 참여한 AOM(Alliance for Open Media)에서 개발한 포맷으로 AV1 코덱을 기반으로 만든 형식입니다. 크기는 역시 JPG의 50%밖에 되지 않고 처음으로 HDR을 지원하는 이미지 형식입니다.

Visual examples



그림 3. 사진 예시

| 원문                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                    | 번역                                                                                                                                                                                                                                                                                                                                                                                              |
|---------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------|-------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------|
| <p>Shown below is one original source image from the Kodak dataset and the corresponding result with JPEG 444 @ 20,429 bytes and with AVIF 444 @ 19,788 bytes.</p> <p>The JPEG encode shows very obvious blocking artifacts in the sky, in the pond as well as on the roof.</p> <p>The AVIF encode is much better, with less blocking artifacts, although there is some blurriness and loss of texture on the roof.</p> <p>It is still a remarkable result, given the compression factor of around 59x (original image has dimensions 768x512, thus requiring 768x512x3 bytes compared to the 20k bytes of the compressed image).</p> | <p>아래에는 Kodak 데이터셋에서 가져온 원본 이미지 하나와 해당 이미지를 JPEG 444 @ 20,429 바이트와 AVIF 444 @ 19,788 바이트로 인코딩한 결과가 나와 있습니다.</p> <p>JPEG 인코딩 결과는 하늘, 연못, 지붕에서 매우 뚜렷한 blocking artifacts*를 보여줍니다.</p> <p>AVIF 인코딩 결과는 blocking artifacts*가 적어서 훨씬 더 좋지만, 지붕 부분에서 일부 흐릿함과 질감의 손실이 있습니다.</p> <p>그렇다고 해도, 이는 59배의 압축 비율을 가지는 것을 감안하면 상당한 결과입니다. (원래 이미지의 크기는 768x512로, 768x512x3바이트가 필요한 반면, 압축된 이미지는 20K바이트입니다.</p> |

\* 압축 과정에서 발생하는 블록 모양 현상 왜곡

원본 사진을 비슷한 용량으로 JPG와 AVIF로 압축했을 때 확연한 품질 차이를 보실 수 있습니다. 더 다양한 예시들을 확인하시려면 원문을 참고하시면 되겠습니다.

## Image quality assessment

우리의 눈에도 확연히 차이가 보일 만큼 원본과의 품질 차이가 심한 경우가 아니더라도 품질 차이를 판단하기 위한 객관적인 기준이 있어야 합니다. 마지막으로 원문에서 사용한 품질 측정법에 대해 간단하게 설명하고자 합니다.

첫 번째는 **BD rate(Bjontegaard-Delta rate)**입니다. 이 품질 측정법은 같은 bit rate에서의 quality 차이, 같은 quality에서의 bit rate 차이를 비교하는 방식입니다. 이 방식에서 말하는 quality란 PSNR(Peak Signal-to-Noise Ratio), 즉 최대 신호 잡음 비와 비트율을 측정해 품질을 비교합니다.

두 번째는 **SSIM(Structural Similarity Index Map)**입니다. SSIM은 PSNR과 다르게 수치적인 예러가 아닌 인간의 시각적 화질 차이를 평가하기 위해 고안된 방법으로 구조적 유사도를 의미하며 휘도(Luminance), 대비(Contrast), 구조(Structural) 3가지 항목을 수식에 대입해 값을 산출한 후 해당 픽셀들의 평균으로 최종 품질 점수를 얻는 방법입니다. SSIM에서 더 발전한 MS-SSIM(Multi-scale SSIM)은 정보의 중요도에 따라 픽셀들의 평균치가 아닌 가중치를 부여해 점수를 산정하는 방법입니다.

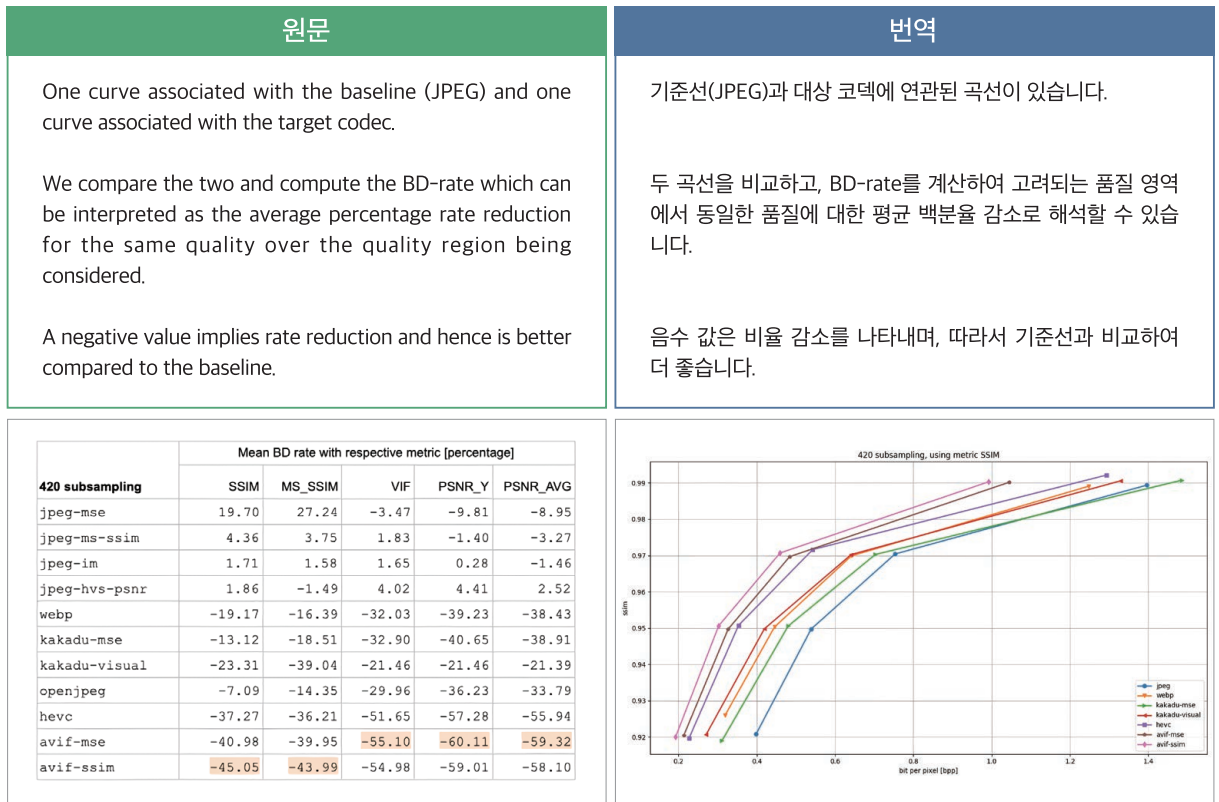


그림 4. 품질비교표

그림 5. 품질비교 곡선

위 원문 내용에서 볼 수 있듯이 AVIF가 JPG보다 동일한 품질을 내기 위해 적은 비트를 사용하고 있는 것을 확인할 수 있습니다. 더 다양한 예시들을 확인하시려면 원문을 참고하시면 되겠습니다.

## AVIF support and next steps

## 원문

At Netflix, we are also working on HDR images for the UI and are planning to use AVIF for encoding these HDR image assets.

## 번역

Netflix에서는 UI를 위한 HDR 이미지를 제작하고 이러한 HDR 이미지를 인코딩하는 데 AVIF를 사용할 계획입니다.



그림 6. PT 토론

아직 AVIF는 JPG에 비해 호환성이 떨어지고 디코딩 시간이 오래 걸리는 등 단점들이 있습니다. 하지만 Google Chrome 등 다양한 브라우저에서 AVIF 이미지 지원을 시작했고, 기사 제목과 마찬가지로 차세대 이미지 파일로의 발전 가능성에 대해서 관심을 가져볼 필요가 있는 것 같습니다.

### 예측 모델링으로 Netflix의 콘텐츠 품질관리 최적화

## 원본 정보

|      |                                                                                                                                                                                                                                               |                                                                                       |
|------|-----------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------|---------------------------------------------------------------------------------------|
| 제목   | Optimizing Content Quality Control at Netflix with Predictive Modeling<br>(예측 모델링으로 Netflix의 콘텐츠 품질관리 최적화)                                                                                                                                    |  |
| Link | <a href="https://netflixtechblog.com/optimizing-content-quality-control-at-netflix-with-predictive-modeling-712281658ab9">https://netflixtechblog.com/optimizing-content-quality-control-at-netflix-with-predictive-modeling-712281658ab9</a> |                                                                                       |

그림 7. 원문 QR-Code

넷플릭스가 지난해 말 기준으로 글로벌 가입자 수가 2억 3,100만 명에 육박했다고 합니다. 이렇게 많은 전 세계 회원들이 수많은 콘텐츠를 다양한 언어를 통해 스트리밍하는 환경이기 때문에, 고품질 콘텐츠 경험을 목표로 하는 넷플릭스는 QC 고도화 및 최적화에 큰 중점을 두고 있습니다. 이러한 배경을 토대로 예측 모델링을 활용한 QC Operation Process가 탄생했다고 볼 수 있겠습니다.

이 글을 읽으시는 분들은 모두 잘 알고 있겠지만 QC(Quality Control)란 품질관리를 지칭하는 용어로 다양한 분야에 통용되어서 사용하고 있습니다. 콘텐츠 시장에서도 QC는 매우 중요한 절차이며 시청자들에게 최종 콘텐츠를 배포하기 이전에 정해진 영상 및 음향 규격에 맞게 콘텐츠가 제작되었는지, 콘텐츠의 품질 및 일관성에 문제는 없는지 검토하는 QC를 수행하게 됩니다. 넷플릭스는 콘텐츠 제작에 요구되는 각종 사항을 Partner Help Center라는 사이트에 모두 공개하고 있는데, 필자 또한 해당 사이트의 Quality Control 자료를 많이 참고하면서 넷플릭스의 QC에 대해 자세히 들여다볼 수 있었습니다.



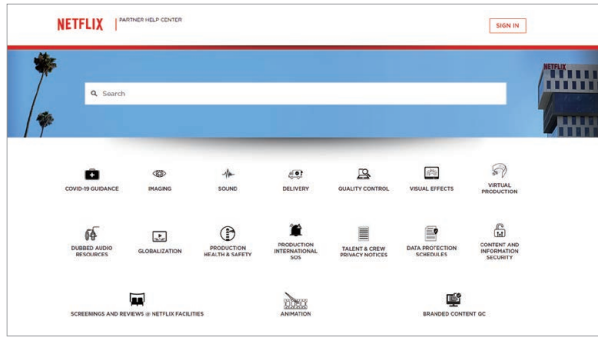


그림 8. Netflix Partner Help Center

/ 출처 : URL : <https://partnerhelp.netflixstudios.com/hc/en-us>

예측 모델링에 대한 설명을 하기 전에 넷플릭스의 QC Operations 절차가 무엇인지에 대해서 먼저 살펴 보겠습니다. 넷플릭스는 QC를 Production QC, QC Operations, Branded Content QC, Localization QC 총 4가지 타입으로 분류하고 있습니다. 이번 기사에서 예측 모델링을 적용한 사례는 QC Operations 타입입니다. QC Operations는 넷플릭스 플랫폼에 배포된 모든 콘텐츠에 대해서 스트리밍에 지장을 줄 수 있는 문제를 식별하는 QC입니다. 주로 잘못된 콘텐츠가 배포되지는 않았는지, 번역 문제나 오타는 없는지, 비디오나 오디오 품질의 결함은 없는지에 대해 초점을 두고 QC를 수행한다고 설명하고 있습니다.

아래 그림에는 QC Operations 절차가 나와 있습니다. Media Asset이 업로드되면 먼저 자동 QC(Auto QC/Inspection as a service)를 수행합니다. 자동 QC는 인코딩 프로세스 전후에 수행되며 콘텐츠가 넷플릭스의 기본 규격을 충족하는지, 지원 가능한 코덱으로 제공되었으며 기술적 오류가 없는지를 점검합니다. 자동 QC 이후에는 수동 QC(Manual QC)를 수행하여 사람의 눈으로 쉽게 감지할 수 있는 문제들을 확인하게 됩니다. 수동 QC는 두 가지 방식으로 수행하게 되는데 예상 인기도가 낮은 콘텐츠는 전체 RT 중에서 몇 부분만 검사하는 Spot QC를 수행하고, 넷플릭스 오리지널 콘텐츠라든지 주목할 만한 콘텐츠는 전체 RT를 체크하는 Full QC를 수행합니다.

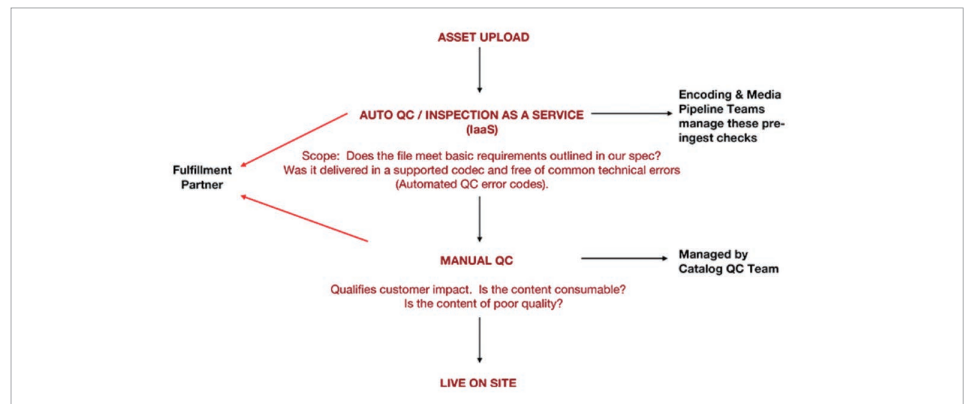


그림 9. QC Operations / 출처 : Netflix Partner Help Center

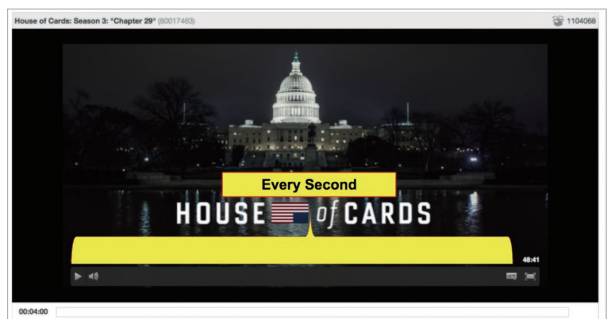
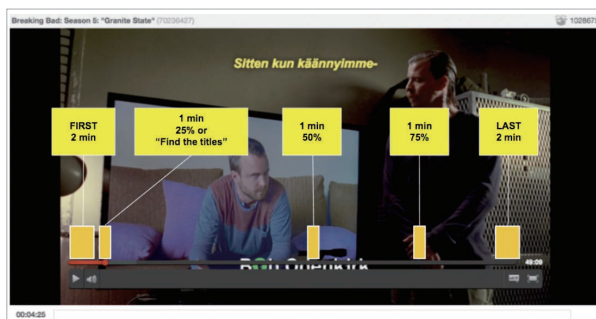


그림 10, 11. Spot QC & Full QC / 출처 : Netflix Partner Help Center

넷플릭스 플랫폼은 전 세계적으로 빠르게 확장되고 있고, 각국의 무수히 많은 제작사에 의해 다양한 콘텐츠를 받는 상황이기 때문에 콘텐츠 품질 관리에 더욱 민감할 수밖에 없습니다. 이러한 상황에서 넷플릭스는 머신러닝 기반의 예측 모델링을 활용한 QC Operations를 도입하게 되었습니다. 아래 그림은 예측 모델링이 적용된 QC Operations 절차입니다.

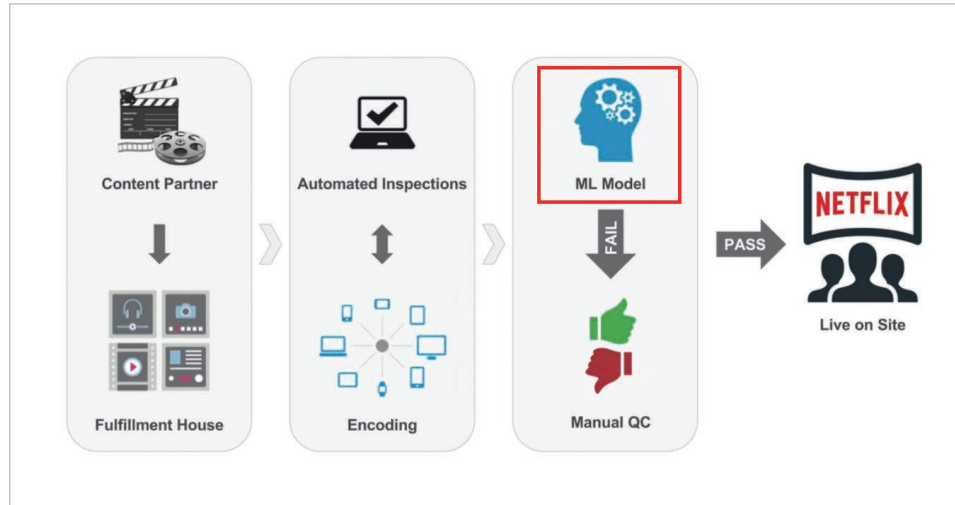


그림 12. Predictive Quality Control / 출처 : Netflix Tech Blog

기존 절차와의 차이점은 QC에 실패할 것으로 예상되는 콘텐츠만 예측 모델링을 통해서 수동 QC로 보내진다는 점입니다. 예측 모델링은 과거의 수동 QC 결과 데이터를 토대로 학습한 머신러닝 기반의 모델입니다. 이를 통해 결함이 없는 콘텐츠에 QC를 수행하는 노력과 비용을 줄임과 동시에 Spot QC에서 놓칠 수 있는 어려운 품질 문제 발견에 리소스를 재할당해서 QC를 강화하는 이점을 가져올 수 있습니다.

예측 모델링의 기반인 머신러닝에 대해서 조금 더 살펴보면, 데이터를 학습할 시 나타날 수 있는 문제점 중 하나는 데이터 불균형입니다. 말 그대로 정상 범위의 데이터와 이상 범위의 데이터가 현저하게 차이 나는 데이터 분포를 불균형 데이터라고 이해하시면 되겠습니다. 금융 사기 데이터, 의료 진단 데이터의 경우 이러한 불균형 데이터가 나타나게 되는데 넷플릭스의 QC 데이터 또한 불균형 데이터 분포를 이루고 있습니다. QC에 문제가 있는 데이터가 QC를 통과한 정상 데이터에 비해서 현저하게 적다는 점 때문에 예측 모델링을 구축할 때 데이터 불균형 문제를 해결해야 합니다. 보통 머신러닝은 이상이 있는 데이터가 타겟이기 때문에 이상이 있는 데이터를 정확히 분류하는 것이 중요하기 때문입니다.

다음 그림은 불균형 데이터 분포에서 발생하는 문제점입니다. 파란색은 QC 정상 데이터이고, 빨간색은 QC 실패 데이터, 회색은 테스트 데이터입니다. 이상적인 분류 경계선이 형성되는 것이 가장 좋겠지만 파란색과 빨간색 데이터만 알고 있는 상황에서는 분류 경계선이 그림과 같이 치우쳐서 형성됩니다. 즉 빨간색 분류 경계선 왼쪽에 있는 회색 테스트 데이터들은 실제로는 QC 실패 데이터이지만 정상 데이터로 오분류 될 수 있기 때문에 정확한 데이터를 찾아 내지 못한다는 문제점이 있습니다.

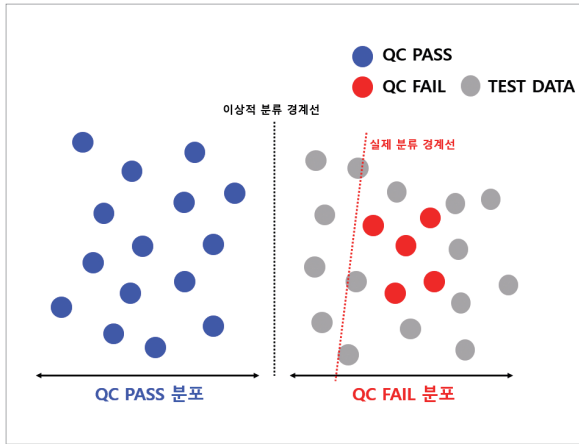


그림 13. 데이터 불균형 문제점

이러한 문제점을 해결하기 위해 크게 Data-level 방법과 Algorithm-level 방법이 연구되고 있는데 넷플릭스에서는 Algorithm-level 기법의 하나인 비용민감학습 (Cost-sensitive learning)을 활용하여 데이터 불균형 문제를 해결했다고 설명합니다. 비용민감학습을 쉽게 설명하자면, 타겟으로 하는 소수 데이터에 가중치를 줘서 데이터 불균형 환경에서도 QC 실패 데이터를 정확히 분류할 수 있도록 모델링하는 학습 기법입니다. 아래 그림을 보시면 좀 더 직관적으로 확인하실 수 있을 것입니다.

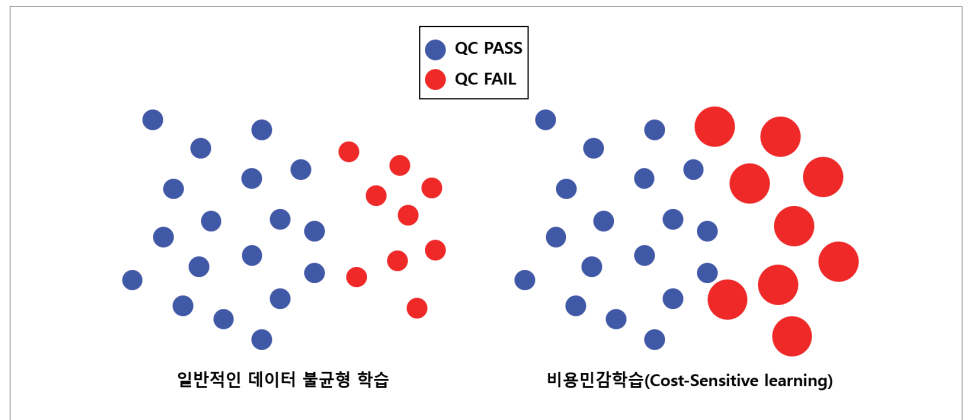


그림 14. 비용민감학습(Cost-sensitive learning)

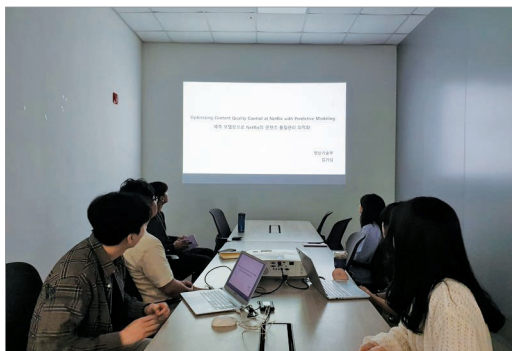


그림 15. PT 토론

이번 넷플릭스 기사를 들여다보면서 가장 인상 깊었던 점은 QC와 관련해서 메뉴얼화가 자세하게 되어있다는 부분입니다. 넷플릭스 오리지널 콘텐츠뿐만 아니라 각국에서 다양한 콘텐츠를 납품받는 환경이다 보니, 납품 시 지켜야 할 기술적 규격을 정말 자세하게 명시해 놓았고 QC Error Code 또한 정리되어 있어서 철저한 품질관리를 수행하고 있다고 느꼈습니다. 때문에 우리 방송사에서도 QC뿐만 아니라 제작 전반에 관련된 자세한 메뉴얼화 및 관리가 더 필요하지 않을까 하는 생각을 가져볼 수 있었습니다.

머신러닝을 현업에 적용하기에는 실제 방송사 환경을 고려할 때 당장 적용하기는 힘들 수 있지만, 머신러닝 기법이 방송기술에 적극적으로 활용되고 있는 시점에서 항상 동향을 주시하고 적용 가능성을 열어두고 있어야 할 것 같습니다. 📺

## 참고문헌

- KIMST 논문(2021). 불균형데이터의 비용민감학습을 통한 국방분야 이미지 연구 성능 향상에 관한 연구
- 고려대학교 DMQA 연구실(2021). Handling imbalanced datasets