



『넷플릭스 기술 연구회』 1부

EBS 기술인협회 스터디
『넷플릭스 기술 연구회』

EBS 기술인협회 스터디 지원 사업 소개

現 EBS 기술인협회(제29대 협회장 김동신)의 정책국은 기술인역량 강화 및 친목 도모를 위해서 기술 스터디 지원 사업을 운영하고 있습니다. 작년인 2022년도 스터디 지원 사업으로 편집부 중심의 ‘파이널컷 프로 연구회’와 NPS, 송출 등의 업무를 담당했던 네트워크기술부 중심의 ‘업무자동화를 위한 RPA 연구회’를 운영한 바 있습니다. 결과적으로 스터디 지원 취지가 잘 구현되어 시사점이 있는 결과가 도출이 됐고, 결과를 사내에 적극 공유하고 협회의 범위를 넘어서서도 활발한 논의가 진행된 바 있습니다.

이에 EBS 기술인협회에서는 올해 역시 스터디 지원 사업을 확대하기로 하였고, 3년 차 미만 주니어 직원이 중심이 되어 음향 제작기술을 원론적으로 다뤄보는 ‘EBS 음향 연구회’, 방송 IT 전문가 중심의 ‘구글 애널리틱스 GA4 연구회’ 그리고 이렇게 ‘방송과기술’을 통해서 소개를 드리려고 하는 ‘넷플릭스 기술 연구회’를 운영하고 있습니다.

『넷플릭스 기술 연구회』 스터디 소개

이미 거대한 규모의 서비스를 운영하는 테크기업이나 스타트업, 특히 기술 스타트업의 경우 내부의 기업문화 홍보를 위해서 S/W, AI, 인프라 등에 대한 기술블로그, 디자이너들을 위한 디자인블로그 등을 운영하는 경우가 많습니다. 실제로 구글이나 아마존 AWS 코리아, Meta, 네이버 등 기술을 선도하는 기업에서부터 획기적인 서비스로 급성장하는 스타트업들이 기술블로그를 운영하고 있습니다.

넷플릭스 역시 오래전부터 ‘미디엄’이라는 블로그 서비스를 통해서 기술블로그를 운영해 왔으며, 지금도 꾸준하게 포스팅을 올리면서 기술을 적극적으로 공유하고 홍보하고 있습니다.



그림 1. EBS 기술인협회 2022년도 스터디 지원사업 - RPA 연구회 결과 보고

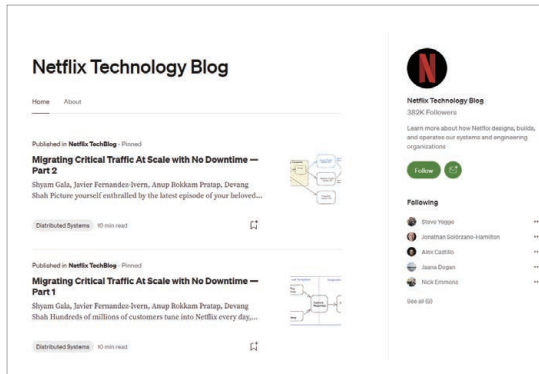


그림 2. 넷플릭스 기술블로그 (netflixtechblog.medium.com)

저희 넷플릭스 기술연구회는 한 협회원이 미디어 현황 조사를 위해 넷플릭스 기술블로그를 연구하던 도중에 ‘이 좋은 내용을 번역해서 협회원들과 공유하면 어떨까?’라는 것에서 시작하게 됐습니다. 취지와 아이디어가 좋았기에 시작은 쉬웠습니다만 그 내용의 효과적인 전달 방법론에 대해서는 많은 고민이 있었습니다. 결국 저희가 선택한 방법은 오픈소스 웹 기술들을 활용해서 내용을 관리·공유할 수 있는 웹서비스를 제작하는 것이었고, 결과적으로 내용 관리와 결과 공유 두 마리 토끼를 한 번에 잡을 수 있었습니다.

Netflix Techblog Article List

번호	제목	번역	작성일	검정	상태
455	What is an A/B Test?	번역	2023-04-07	2023-07-28	검정
279	Using Machine Learning to Improve Streaming Quality at Netflix	번역	2023-04-07	2023-05-25	검정
490	Machine Learning for Fraud Detection in Streaming Services	번역	2023-04-07	2023-05-25	검정
3	Four Reasons We Choose Amazon's Cloud as Our Computing Platform	번역	2023-04-10	2023-04-27	검정
218	Netflix Recommendations: Beyond the 5 stars (Part 1)	번역	2023-04-28	2023-06-16	검정
2	Why we use and contribute to open source software	번역	2023-04-07	2023-05-25	검정
240	How Data Science Helps Power Worldwide Delivery of Netflix Content	번역	2023-05-25		검정
246	Content Popularity for Open Connect	번역	2023-05-25		검정
288	Engineers Making Movies (AKA Open Source Test Content)	번역	2023-04-11	2023-05-25	검정
312	Enhancing the Netflix UI Experience with HDR	번역	2023-05-30	2023-06-16	검정
380	AVIF for Next-Generation Image Coding	번역	2023-04-07	2023-04-27	검정
165	Optimizing Content Quality Control at Netflix with Predictive Modeling	번역	2023-04-07	2023-05-25	검정
166	Per-Title Encode Optimization	번역	2023-06-12		검정
1	Netflix Tech Blog				
4	5 Lessons We've Learned Using AWS				

그림 3. 결과 공유 웹서비스 제작 - 리스트 관리

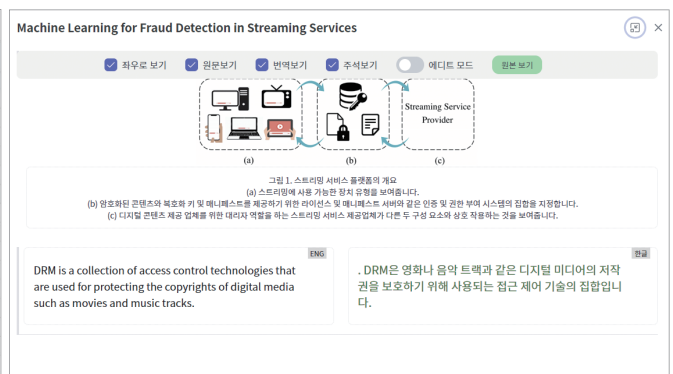


그림 4. 결과 공유 웹서비스 제작 - 포스트 뷰어

사실 본인에게 익숙하지 않은 언어로 표현된 내용이 번역된다는 것만으로도 관심 있던 내용에 대한 접근성은 크게 향상됩니다. 물론 요즘은 웹서비스 차원에서 자동 번역이 꽤 훌륭하게 처리되는 편이지만, 그 기술에 대한 경험이 있는 사람이 번역을 수행하고, 그 내용을 우리 현실에 비춰서 한 번 더 고민해서 그 결과를 공유한다면 그야말로 의미 있는 결과가 도출됩니다. 당초 이 스터디를 운영할 때는 본인의 업무 외에 투자하는 과외활동이기에 번역만 올리고, 그 분야에 관심 있는 직원분들이 보고 각 개인의 insight를 얻어가면 다행이라고 생각했습니다. 하지만 스터디원으로 참여하신 신입사원분들께서 적극적이고 능동적인 의지를 보여서 번역본과 함께 시사점에 대한 PPT를 따로 제작해 오시는 결과를 보여주셨습니다.

그리고 그 결과를 놓고 매월 2회 정도씩 토론의 장이 벌어지는데, 이 내용이 우리만 알고 지나가기에는 아까운 느낌이 들었고, 이렇게 ‘방송과기술’을 구독하시는 많은 분들과 공유하고 싶어서 연재를 시작하게 됐습니다.

계획대로 진행된다면 앞으로 6회의 걸쳐서 저희가 번역하고 토론했던 내용 중 일부를 선정하여 공유해 드릴 예정입니다. 각 회차에는 2개의 게시물과 번역본, 그리고 저희가 고민했던 내용을 함께 올려서 모두가 함께 긍정적인 고민을 하시고, 발전할 수 있는 고무적인 결과가 있기를 희망합니다. 회차별로 필자가 다를 수 있어서, 내용 구성이나 표현 또한 매회 다를 수 있다는 점도 미리 양해 부탁드립니다.

넷플릭스의 오픈소스 정책

원본 정보

제목	Why we use and contribute to open source software (넷플릭스가 오픈소스 소프트웨어를 사용하고 기여하는 이유)
Link	https://netflixtechblog.com/why-we-use-and-contribute-to-open-source-software-1faa77c2e5c4



그림 5. 원문 QR-Code

포스팅 개요

이 포스팅을 선정한 이유는 이번 담당자가 소프트웨어 분야 특히 Open Source에 큰 관심이 있었던 것을 첫 번째로 꼽을 수 있겠습니다. 넷플릭스는 2010년부터 기술블로그를 운영해 왔는데 지금까지 500개가 넘는 기사가 쌓여있습니다. 이 중 'Open Source'로 검색이 되는 기사는 8개가 있는데, 그중 넷플릭스가 Open Source를 보는 기본적인 시각을 잘 알 수 있는 기사가 바로 첫 번째 기사인 「Why we use and contribute to open source software」입니다.

넷플릭스의 시작은 알려진 바와 같이, 비디오 대여 사업으로, 이 비디오 유통 사업이 DVD의 등장과 동시에 폭발적으로 성장했고, 그 역사가 지금의 스트리밍 사업으로 연결되어 있습니다. 당시에 넷플릭스는 처해 있는 환경에서 혁신적인 시스템을 도입하곤 했었습니다. 대여물의 회수가 기대만큼 원활하게 되지 않자 연체료를 도입했고, 연체료가 있음에도 불구하고 회수가 잘 안 되자, DVD 3개를 무제한 대여하되, 기존의 대여물을 반납하지 않으면 새로운 대여를 할 수 없는 파괴적인 정책을 도입했었습니다. 비디오가 내구성 좋은 DVD로 변화하자 직접 배달하는 서비스를 도입했고, 지금은 사람이 아닌 IT 인프라가 원하는 미디어를 TV까지 배달해 줍니다. 이를 위해서 일반인은 알 수 없는 Back-end에서 많은 작업이 수행되고 있으며, 넷플릭스는 이 시스템을 위해서 다양한 오픈소스를 활용하고 그 결과물을 다시 기여하는, 오픈소스 생태계를 정확하게 이해하고 있는 모습으로 기업을 운영하고 있습니다.

이 포스트를 작성한 인물은 당시 넷플릭스의 기술 부사장으로 처음 던지는 말이 이것입니다.

〇〇 Netflix is a technology company. 〇〇

이 말은 넷플릭스가 IT 기술회사라는 것을 설명하는 것이 아닙니다. 짧은 한 줄이지만, 개인적으로 'Digital Transformation'에 대한 많은 과정과 결과가 담겨있는 말이라고 읽혔으며, 실제로 넷플릭스는 '비디오 대여 → DVD 배달 → 스트리밍 서비스'의 과정으로 Digital Transformation에 성공한 기업입니다. '우리 회사에서 기술은 중요하다.'라는 말은 누구나 쉽게 할 수 있지만 이렇게 서비스를 운영하는데 있어서 적극적이고 도전적인 자세로 기술을 대하는 기업은 드뭅니다. (특히 태생이 기술 기반 회사가 아니었음에도 불구하고 말입니다!) 다음은 넷플릭스가 오픈소스를 활용하는 이유를 명확하게 드러낸 단락입니다.

원문	번역
Our budget, measured in dollars, time, people, and energy, is limited and we must therefore focus our technology development efforts on that streaming video software that clearly differentiates Netflix and creates delight for our customers. These limits require that we stand on the shoulders of giants* who have solved technology challenges shared in common by all companies that operate at Internet scale.	비용, 시간, 인력 그리고 에너지와 같은 넷플릭스의 자원은 제한적이기 때문에, 우리는 우리의 고객에게 넷플릭스만의 확실한 차별점과 즐거움을 제공하는 비디오 스트리밍 기술 개발 노력에 집중해야만 합니다. 이러한 한계(자원의 제한 속에서 기술개발에 매진해야 하는 상황)는 우리가 '인터넷 규모로 운영되는 모든 기업이 공통으로 공유하는 기술 과제를 해결한 거인'의 어깨에 서 있도록 합니다.

* Standing on the Shoulders of Giants : 아이작 뉴턴이 대단한 발견을 한 후에 '자신보다 앞서 살았던 훌륭한 과학자들이 그들이 이루어 놓은 것(아래) 위에서 더 멀리 볼 수 있었다.'라며 겸손을 표한 것.

단순한 코드의 공유 차원이 아닌 생태계로서 오픈소스의 장점은 너무나 명확합니다. 낮은 TCO, 다양한 커뮤니티, 표준 및 보안에 대한 신뢰성 등. 하지만 이는 잘 구축된 IT 거버넌스 아래에서 그 진가를 발휘하게 되며, 경직된 레거시 의사결정 구조를 그대로 가지고 있는 회사에서 오픈소스(클라우드 등 IT 솔루션을 의미하는 것)를 턴키 방식으로 도입하는 것은 오히려 해외 기술의존도를 높이고 기술변화에 능동적인 대응이 힘들어질 수도 있는 결과를 불러옵니다.



그림 6. 내부 보고용 PT 및 토론

이번에 소개한 포스팅은 넷플릭스 스터디 시스템 구축을 테스트하는 차원에서 짧고, 의미가 있는 포스팅을 선택했습니다. 따라서 포스팅의 내용에 집중하는 것보다는 그 의미를 생각해 보는 시간이 길었는데, 오히려 방송사의 IT 거버넌스 구축이라는 차원에서 고민은 더 깊어지는 시간이었습니다.

머신러닝을 통한 스트리밍 품질 향상

넷플릭스는 맞춤형 콘텐츠 추천뿐만 아니라 콘텐츠 제작, 비디오 품질 향상, 서비스 부정 사용 모니터링, 내부 운영 및 개선 등 다양한 분야에서 머신러닝 기술을 활용하고 있습니다. 넷플릭스 기술 블로그에서도 다양한 머신러닝 활용 사례를 소개하고 있는데요, 머신러닝과 관련된 첫 주제로 머신러닝을 활용한 스트리밍 품질 향상 방안을 소개해 드리고자 합니다.

원본 정보

제목	Using Machine Learning to Improve Streaming Quality at Netflix (스트리밍 품질 향상을 위한 머신러닝 기법의 활용)	
Link	https://netflixtechblog.medium.com/using-machine-learning-to-improve-streaming-quality-at-netflix-9651263ef09f	

그림 7. 원문 QR-Code

스트리밍 품질 개선 배경

넷플릭스는 전 세계 1억 1,700만 명 이상의 회원에게 스트리밍을 제공하고 있습니다. 글로벌 시청자에게 더 나은 시청 경험을 제공하기 위해서는 서버 기술과 함께 스트리밍 알고리즘에 대한 고려가 필요합니다. OTT 서비스의 다양한 시청 행동 특성과 사용자별 환경이 상이한 점을 고려할 때 스트리밍 품질 개선을 위하여 일반화된 해결책을 적용하는 것은 한계가 발생할 수밖에 없습니다.

OTT 서비스 환경		다양한 서비스 환경		다양한 시청자별 특성, 네트워크 환경 및 기기의 성능	
		최적의 대응 방안		모든 세션에 대해 네트워크와 기기 상태, 비디오 품질과 같은 사용자 경험을 수집, 통계 모델링과 기계 학습을 활용	
시청 및 검색 특성		변동성이 큰 이동통신망 특성		네트워크 혼잡 상황	
스마트TV, 모바일 다양한 시청/브라우저 특성		고정된 대역폭 네트워크 보다 변동성이 크고 불안정한 이동통신망의 특성		일부 서비스 지역에서 발행할 수 있는 높은 네트워크 혼잡 상황	
				다양한 하드웨어 사양	
				하드웨어 사양에 따른 상이한 인터넷 용량과 인터넷 연결 신뢰성	

사용자에게 고품질 경험을 제공하고 새로운 시장에서 확장하기 위해서는 이러한 다양한 환경에 적응적으로 대응할 수 있어야 합니다. 이를 위하여 넷플릭스에서는 모든 세션에 대해 네트워크와 기기 상태, 비디오 품질과 같은 사용자 경험 측면을 관찰하여 통계 모델링과 기계 학습을 활용하고 있습니다.

Network quality characterization and prediction

네트워크 품질 지표인 네트워크 평균 대역폭, RTT(Round Trip Time)와 함께 안정성과 예측 가능성을 고려하여야 최적의 비디오 스트리밍 품질을 확보할 수 있습니다. 이러한 네트워크 품질을 특성화하고 예측, 분석하는 것은 초기 비디오 품질을 결정하거나 재생 중 비디오 품질을 조정하는 데 도움을 줄 수 있습니다.

다음은 실제 시청 세션에서 측정된 몇 가지 네트워크 처리량 예시입니다. 노이즈가 많고 다양한 범위 내에서 네트워크 처리량이 변동되는 특성을 갖습니다.

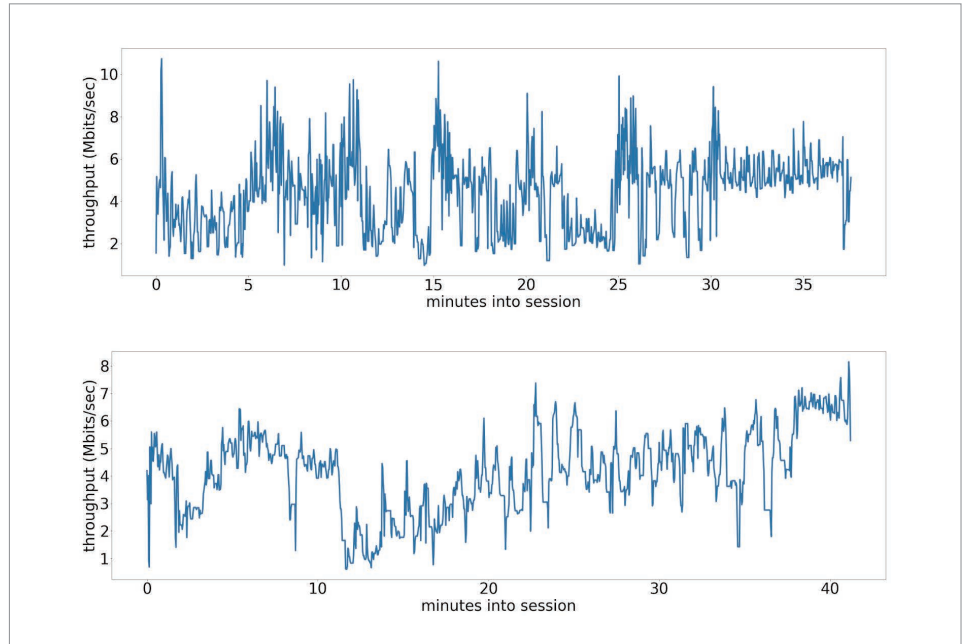


그림 8. 실 환경에서의 네트워크 Throughput 측정 데이터

네트워크 특성을 분석하여 현재의 네트워크 상태를 바탕으로 이후의 처리량을 예측하거나 기기에 최적화된 트래픽을 제공하기 위해 데이터율을 가변하는 등의 프로세스를 통해 최적의 서비스를 제공할 수 있습니다. 이를 위하여 넷플릭스에서는 대규모 추적 데이터를 수집, 분석하고 시간적 패턴 인식과 다양한 맥락적 지표를 결합하여 정확한 예측을 수행할 수 있는 AI 모델을 개발하여 적용하고 있습니다.

이번 포스팅에서는 스트리밍 품질 향상을 위해 머신러닝이 적용되고 있는 Video Quality adaptation, Predictive Caching, Device anomaly detection에 대해 간략하게 소개하고자 합니다.

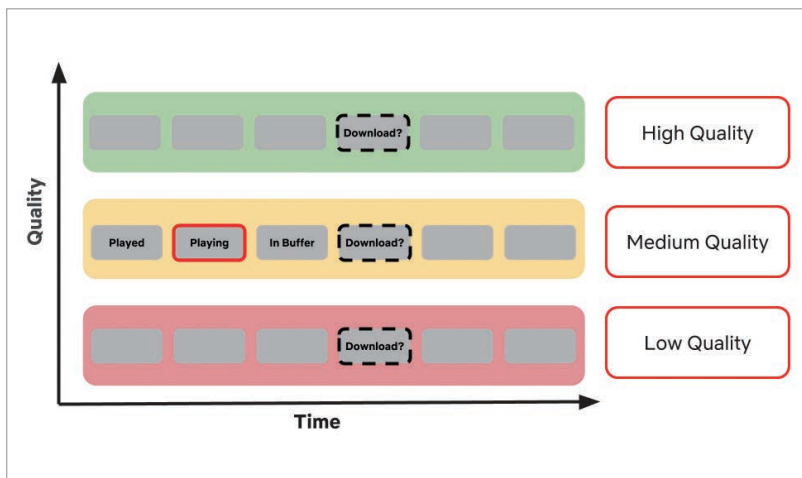


그림 9. 네트워크 품질에 따른 콘텐츠 품질 서비스

Video quality adaptation during playback

넷플릭스 콘텐츠는 다양한 네트워크 환경과 기기의 성능을 지원하기 위하여 여러 가지 비디오 품질로 인코딩됩니다. 여러 해상도와 비트레이트로 인코딩된 콘텐츠를 통해 사용자의 환경이 원활한 경우 고화질 비디오를 제공하고, 연결이 불안정한 경우에는 낮은 해상도의 비디오로 전환하여 버퍼링을 최소화합니다.

적응형 스트리밍 알고리즘은 현재의 네트워크와 기기 조건인 QoE(quality of experience)에 기반하여 재생 중에 스트리밍되는 비디오 품질을 조정하는 역할을 담당합니다. QoE는 비디오 재생을 시작하기까지 기다리는 시간, 사용자가 경험하는 전체적인 비디오 품질, 재생 중에 추가로 비디오를 버퍼에 로드하기 위해 재생이 일시 중단되는 횟수, 재생 중 품질의 인지 가능한 변동량 등과 같은 여러 가지 방법으로 측정됩니다. 이러한 변수들은 Trade-off 관계에 있습니다. 예를 들어, 고화질 스트리밍은 재버퍼링의 위험을 증가시키거나, 재버퍼링을 줄이기 위해 사전에 다운로드하는 경우 사용자의 대기 시간을 늘릴 수 있습니다. 강화학습을 통해 사용자의 QoE를 분석하고 적응형 스트리밍 알고리즘에 필요한 변수들을 학습하여 최적의 적응형 스트리밍 서비스를 제공할 수 있습니다.



그림 10. Video Quality adaptation 제공 절차

Predictive Cashing

통계적 모델은 사용자가 재생할 것으로 예측하여 사용자가 재생하기 전에 장치에 캐싱하는 방식으로 스트리밍 품질을 향상하는 데 활용됩니다. 예를 들면, 특정 시리즈를 시청하는 경우 다음 에피소드를 재생할 확률이 매우 높습니다.



그림 11. Predictive Caching 적용 절차

넷플릭스에서는 사용자의 시청 기록, 상호작용 및 기타 문맥 변수를 바탕으로 지도학습을 수행하여 영상을 캐싱, 대기 시간을 최소화하고 스트리밍 품질을 향상하고 있습니다.

Device anomaly detection

넷플릭스는 노트북부터 태블릿, 스마트 TV, 모바일 폰, 스트리밍 스틱까지 1,000가지 이상의 다양한 종류의 디바이스에서 운영됩니다. 신규 디바이스와 기존 장치의 펌웨어 업데이트, 넷플릭스 애플리케이션 업데이트 등에 의해 앱 실행, 재생 등에 영향을 줄 수 있습니다.

기존에는 이러한 문제를 파악하기 위해 수작업을 진행했습니다. 이 과정에서 예측 민감도를 낮추는 경우 문제를 놓치는 경우가 발생하였으며, 엄격한 기준을 적용한 경우에는 불필요한 수작업이 발생하기도 했습니다.

