



데이터 마이닝으로 게임 이탈 예측하기

(Predicting Churn: Data-Mining Your Game)

작성자: 드미트리 노즈닌 (Dmitry Nozhnin)

작성일: 2012 년 5 월 17 일

모든 온라인 서비스와 게임의 슬픈 현실이라면? 게임 플레이의 가장 초반에 유저 이탈이 제일 심하게 일어난다는 점이다. 이 문제에 대한 연구는 인터페이스의 사용성과 단순성, 무료 체험 기회, 학습 곡선, 튜토리얼의 퀄리티 등과 연관된 다양한 가설에 따라 이미 여러 방면으로 이루어져 왔다. 이러한 요소들이 모두 매우 중요하게 생각되고 있다.

우리는 왜 신규 플레이어가 일찍 게임을 그만두는지를 밝히고, 또 어떤 플레이어가 이탈하려 하는지를 예측해 보려고 한다. MMORPG 게임인 <아이온(Aion)>의 사례를 연구 대상으로 삼았지만 결과는 놀랍게도 광범위한 서비스와 게임에 적용할 수 있는 것으로 드러났다. 이 연구가 진행될 당시에 <아이온>은 7 일간 무료로 게임을 체험해보고, 20 레벨까지 올려볼 수 있는 순수한 회원제 게임이었으나, 많은 유저들은 돈을 내야 하는 시점 훨씬 이전에 게임을 그만 두었다. 우리의 연구는 게임 내에서 이런 이탈을 유발하는 도화선이 무엇인지에 관한 것이다.

행동적 측면의 연구에서 가벼운 게임 이용자들은 주의 집중시간이 그리 길지 않음이 나타났다. 이들은 오늘 게임을 접고 나면 내일은 이 게임이 설치돼 있는지, 플레이를 한 적이 있는지조차 기억하지 못할 수도 있다. 우리는 그들이 게임을 떠나는 **즉시** 다시 돌아오게끔 해야 한다.

그러나 어떻게 저녁에 할 일이 있어서 한동안 접속을 못하는 일상적 게임 이용자와, 게임을 접고 떠나려 하는 사람을 구분할 수 있단 말인가? *이상적인* 방법은 *게임 플레이 중*에 게임을 접을 가능성을 예측하는 것이다. - 플레이어가 실제로 게임을 관두겠다고 생각하기도 전에 나타나는 행동들 말이다.

우리의 목표는 더 현실적이었다. **마지막으로 접속한 날** 신규 플레이어의 이탈을 예측하기로 한 것이다. 우리는 이탈을 7일간 게임을 하지 않는 것으로 정의하고, 플레이어가 게임을 이탈했는지 판단하기 위해 일주일을 기다리는 대신 게임을 한 마지막 날 이를 예측한다는 목표를 세웠다. 우리는 미래를 예측하고 싶었던 것이다!

기술적 측면

우리는 방대한 양의 데이터를 보유하고 있었다. 다행히도, <아이온>의 로그 기록 시스템은 한국에서 출시된 게임 중 가장 훌륭했다. 문자 그대로 플레이어의 모든 움직임과 활동을 추적한다. 초반 10 레벨 혹은 10 시간의 게임 플레이의 데이터를 저장했으며, 50 퍼센트 이상의 초반 이탈자들을 포착했다.

32GB 램과 10TB 콜드 스토리지와 3TB 핫 스토리지 RAID10 SAS 유닛을 장착한 두 대의 듀얼 제온(Dual Xeon) E5630 장비가 사용되었다. 양쪽 장비에는 MS SQL 2008R2 를 실행했는데, 하나는 데이터 보관소로, 다른 하나는 MS Analysis Services 용이었다. 표준 Microsoft BI 소프트웨어만 사용했다.

1 단계: 난 다 알고 있어!

나는 게임기획자로서 100 개도 넘는 플레이테스트를 거치며 많은 경험을 축적했기에, 내 전문성이 이탈에 대한 모든 해답을 대답을 내줄 것이라는 자부심이 있었다. 텔레포트 하는 법을 배우는 데 실패한 사람은 게임을 접었다. 첫 번째로 맞닥뜨린 몸에게서 치명타를 입은 사람도 게임을 접었다. "미션"탭을 잊어버리고 다음에 뭘 해야 될지 모르는 사람도 게임을 아마 접을 것이다. <아이온>은 시각적으로 훌륭하고 기술적으로도 뛰어나지만, 신규 플레이어에게 친절한 게임이라고 하기는 힘들었다.

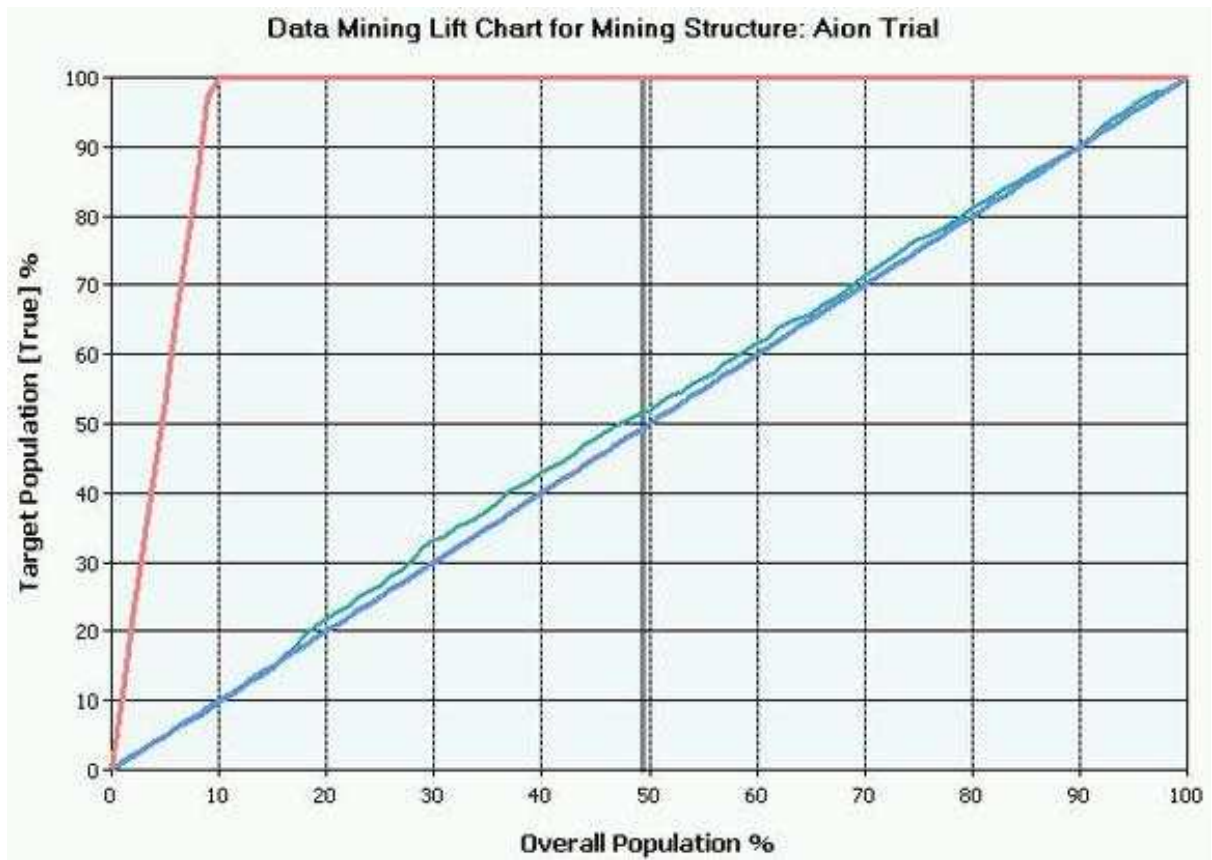
그래서 나는 "일반 플레이어"가 되어 두 종족과 몇몇 직업으로 <아이온>의 무료 체험 기간을 체험하며, 게임 플레이 중 벌어지는 일들을 꼼꼼하게 관찰하고, 이탈의 근본 원인을 설명하는 기초 가정을 세웠다.

- 종족과 직업. 나는 이것이 핵심 요소라고 가정했다. 지원 성향이 강한 사제는 강력한 마법사와는 게임플레이가 다르기 마련이며, 이는 결과적으로 플레이어의 즐거움에도 영향을 주기 때문이다.
- 플레이어가 이노바(Innova)¹의 다른 게임을 해본 적이 있는가? (우리에겐 계정이 하나뿐이었다.)
- 얼마나 많은 캐릭터를 어떤 조합의 종족과 직업으로 만드는가?
- 체험기간 중 레벨당 사망 횟수와 총 사망 횟수
- 다른 플레이어와의 협동 (고레벨과 저레벨)
- 편지 수신과 길드 가입 횟수 (숙련된 플레이어가 키우는 "부캐릭터"라는 신호)
- 퀘스트 완료의 레벨당 횟수와 총 횟수
- 전투 중 사용한 기술의 다양성

이 목록은 인상적이고 세부적이었으며, 게임에서 플레이어가 이탈하게 되는 엄청나게 다양한 이유를 설명하고 있다.

¹ <아이온>의 퍼블리싱 파트너

자 이제 본격적으로 돌입해보자. 첫 번째 가설은 데이터 마이닝 모델이다. 개념은 아주 간단하다. 우리는 **이탈 여부(is leaver)**라는 이름을 가진 불 변수의 값을 하나 예측하는데, 이것은 플레이어가 오늘 게임을 이탈할지, 아니면 남아서 최소한 한동안은 더 게임을 즐길지를 말해주는 것이다.



리프트차트 101: 바닥의 직선은 단순한 무작위 추측이다. 위로 치솟은 선은 뛰어난 이들의 것으로, 미래를 정확히 알고 있는 사람들이다. 이 사이에 나타난 얇은 선이 우리의 데이터 마이닝 모델이다. 선이 수직선에 가까워 질 수록, 미래를 예측하는 힘이 더 강해지는 것이다. 이 차트는 레벨 7 플레이어들의 것이지만, 레벨 2에서 9 레벨도 똑같은 그림이다.

망했다! 우리의 첫 번째 모델은 동전을 던져서 미래를 예측하는 수준을 넘지 못하는 것이었다. 이제 마이닝 스트럭처에 다른 가설을 가져와, 이를 실행해보고 잘 되길 빌어볼 차례다.



조금 나아보이긴 하지만, 아직도 정확도²는 50 퍼센트를 조금 넘는 정도이고, 긍정오류³의 가능성도 28 퍼센트라는, 어마어마한 수준이었다.

² Precision 조건에 맞는 원소가 선택된 비율

정확도와 재현율 101: 정확도가 높아질 수록, 더 많은 진짜 이탈자들을 모델에서 추적할 수 있다. 긍정오류는 이탈자로 예측된 플레이어들이, 실제로는 그렇지 않았을 경우를 말한다.

1 단계 결과: 내 초반 계획이 실패했다. 대재앙이다!

2 단계: 전 완전 초짜예요.

처음 실행한 가장 간단했던 데이터 마이닝 알고리즘은 나이브한 '베이즈(Bayes)'로서, 극도로 인간 친화적이고 이해가 가능한 것이었다. 이것은 가설상의 측정자료는 실제 이탈자들과 연관이 없었음을 보여주었다. 두 번째 방법인 의사결정 트리 모형은 내 아이디어 중 몇몇은 실제로 유용했음을 보여주었지만, 예측 정확성을 최고치로 높이기엔 모자랐다.

데이터 마이닝 알고리즘 101: 나이브 베이즈 분류법(Naive Bayes Classification)의 원리는 데이터셋의 기초 분석과 변수간의 상관관계를 밝히는 데에는 매우 유용하다. 의사결정 트리(Decision Tree)는 데이터셋을 서로 구분되는 부분집합으로 분할해 주므로, 이탈 유저와 행복한 플레이어를 분리하는 데에 유용하다. 신경망(Neural Network)는 기본적으로는 복잡한 변수 관계를 고려하여 더 나은 예측을 해내는 블랙박스이고, 개발자가 그 속을 들여다 볼 수 없다는 단점이 있다.

나는 <아이온>팀과의 브레인 스토밍을 통해 초보 플레이어들에 대해 논의할 수 있는 소중한 기회를 가졌다. 그들이 누구이고, 어떻게 플레이하며, 뚜렷한 특성이 무엇인지에

³ False positives rate 아닌 것을 맞다고 예측할 확률

대해 말이다. 우리는 친구, 친척들이 어떻게 처음 게임에 들어왔으며, 첫 경험이 어땠는지를 돌아보았다.

이 브레인스토밍 세션의 결과로, 초보 플레이(인벤토리 사이즈는 늘렸는가, 부활 거점은 등록했는가, 이동속도 증가 두루마리는 사용했는가?)에 영향을 미치는 게임 내 요소의 목록을 업데이트 할 수 있었으며, 또한 **일반적인 게임 내 활동을** 측정할 수 있는 반짝이는 아이디어도 얻었다.

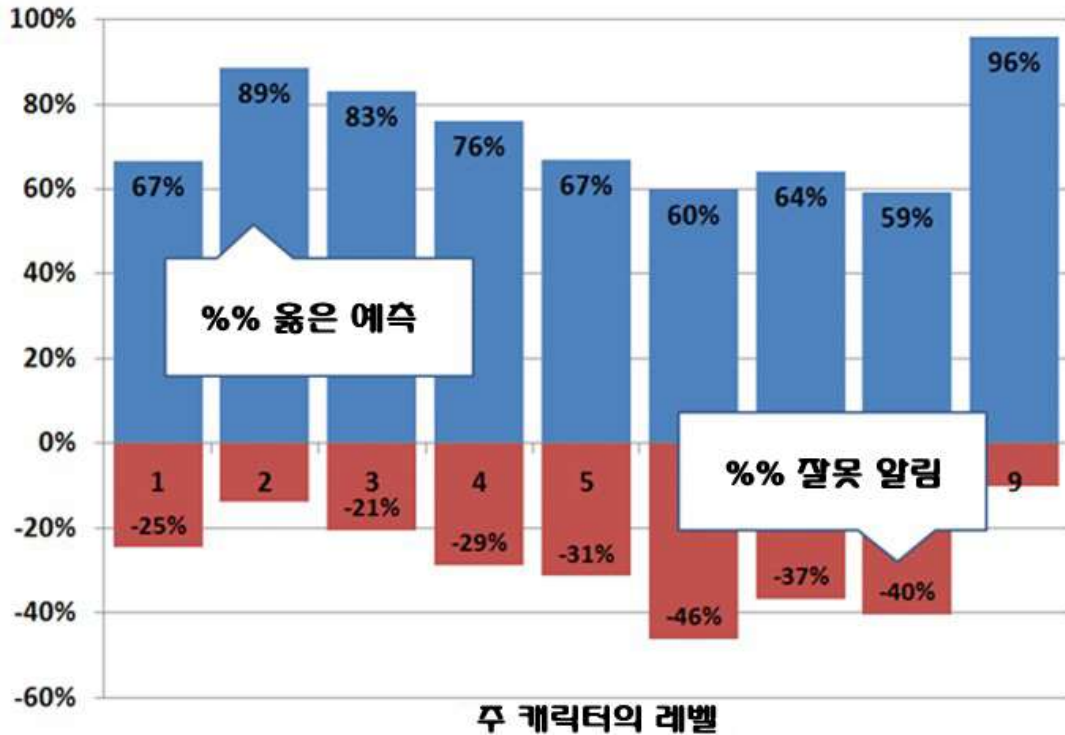
그리고 다음 수치를 이용했다.

- 레벨 당 몹 킬 수
- 레벨 당 완료 퀘스트 수
- 레벨 당 플레이 시간

그 시점에서 ETL(추출, 변환, 로딩을 뜻하는 머리글자-extraction, transformation, loading) 부분을 완전히 뜯어 고쳤고, 우리의 SQL 엔지니어는 로그에 추가된 새 게임 이벤트와 확장성에 초점을 맞춘 SSIS 기반의 게임 로그 처리기를 만들어냈다. 주어진 로그 데이터의 양이 수 기가바이트에 달하였으므로, 새로운 가설을 용이하게 추가하는 기능이 필수적이었다.

새 데이터를 로딩하고 처리하고, 모형을 시험하고 검토하고, 그 결과를 분석하였다. 여기에서는 다른 리프트 차트는 생략하고 정제된 결과만 여기에 올리겠다.

레벨 별 정확성과 재연성



레벨 9 에서의 비정상적으로 높은 정확도는 연구 시점의 게임 자체와 연관되어 있기 때문에(체험기간이었으므로), 그 자료는 무의미하다.

이 단계에서, 우리 모델은 특히 레벨 2와 4 사이에서 그 예측력이 향상됐다. 하지만 레벨 6에서 8까지는 아직 너무 부정확했다. 이런 부정확한 결과는 별로 쓸모가 없다.

의사결정 트리는 일반적 행동 측정이 핵심 측정 요소임을 증명했다. 어떤 면에선 레벨 당 플레이시간, 레벨 당 몹 킬 수, 레벨 당 완료 퀘스트 수가 우리 모델의 예측력 중 핵심을 구성하는 것이었다. 다른 측정치들은 전체 정확도에 기여도가 5 퍼센트 미만이었다. 또한, 의사결정 트리는 무척 간결하여 두세 개의 브랜치만을 가졌기 때문에, 관련된 측정 자료가 부족한 면이 있었다. 왜 이 세 알고리즘이 모두 레벨마다 다른 정확도와 재현율을 보여주었는지는 아직도 나에게 수수께끼로 남아있다.

2 단계 결과: 우리는 특정 게임의 콘텐츠와 관련된 것이 아닌, 일반적인 활동 측정에 있어서 괄목할 만한 성과를 거두었다. 정확도는 아직 만족할 만하지 못했지만, 베이스 분류법을 먼저 쓴 후에 트리를 적용하는 것이 분석에 적합하다는 것을 알아내었다.

3 단계: 올바른 길

데이터 마이닝 결과가 눈에 띄 정도로 향상된 것에 고무되어 나는 3 가지 개발 방향을 잡았다. 1) 더 많은 일반 활동 측정 자료, 2) 더 많은 게임 콘텐츠 별 측정 자료, 3) 더 깊은 마이크로소프트 BI 툴의 이해.

일반 활동에 대한 실험을 통해, 우리의 요책은 다음과 같음을 밝혀내었다:

- 현재 레벨, 이전 레벨, 전체 게임 기간 별로 집계한 플레이 시간
- 분당 몹 사냥 횟수 (현재/이전/전체)
- 일 평균 플레이 시간
- 플레이한 날짜 수
- 결석률 (7 일간의 무료 체험기간동안 접속하지 않은 날)

재현율 증가가 높은 경우는 이들 수치가 관련되어 있었다. (이는 긍정 오류가 적다는 뜻이므로 좋은 소식이다!) 의사결정 트리도 마침내 폭발적으로 마구 가지를 뺀기 시작했다. 또한 서로 달랐던 데이터 마이닝 알고리즘이 모든 레벨에서 통합되는 것을 발견했다. 이는 무작위 결과에서 벗어나 예측 과정이 안정화 되어가고 있다는 좋은 신호였다. 나이브 베이스는 의사결정 트리와 신경망보다 정확도 면에서 10 퍼센트 가량 저조한 결과를 보였다.

새 개인별 측정자료는 실제로 운용하는데 많은 어려움이 따랐다. 자동 공격을 수작업으로 구분하는 데에는 약간의 수학적 지식과, 상위 75% 지점의 계산을 SQL 로 처리하는 등의 작업이 필요했다. 하지만 데이터를 정규화하자, 서로 다른 직업을 비교할 수 있게 되었다.

데이터 마이닝 모델은 이제 단순한 원자료가 아닌, 카테고리로 색인된 데이터가 되었다. 측정자료의 정규화와 색인 작업을 새로이 한 덕분에 전체 정확도는 3 에서 4 퍼센트 가량이 확실하게 향상되었다.

전투 101: 온라인 게임에서, 캐릭터는 기술과 능력치를 통해 싸운다. 자동공격은 가장 기본적이고 수고가 들지 않는 행위이다. 숙련된 플레이어들은 가능한 모든 기술을 사용하므로 그들의 자동 공격률은 줄어들 것이다 -- 게임과 종족의 작동기제가 이를 측정하는데 막대한 영향을 미친다 해도 말이다. <아이온> 에서, 마법사의 중간값은 전사가 70 퍼센트에 육박할 때 5 퍼센트에 그쳤으며, 심지어는 같은 종족 안에서도 표준편차가 높았다.

다음으로 한 일은 [Data Mining with Microsoft SQL Server 2008](#)⁴이라는 책을 읽으며 분석작업에 대한 요령과 기술을 탐색한 것이었다. 책 자체는 의사결정 트리의 미세조정에 유용한 이모저모를 설명하는 것이었으나, 나는 이 책을 통해 정확한 데이터 분할의 중요성에 대해 깨달을 수 있었다.

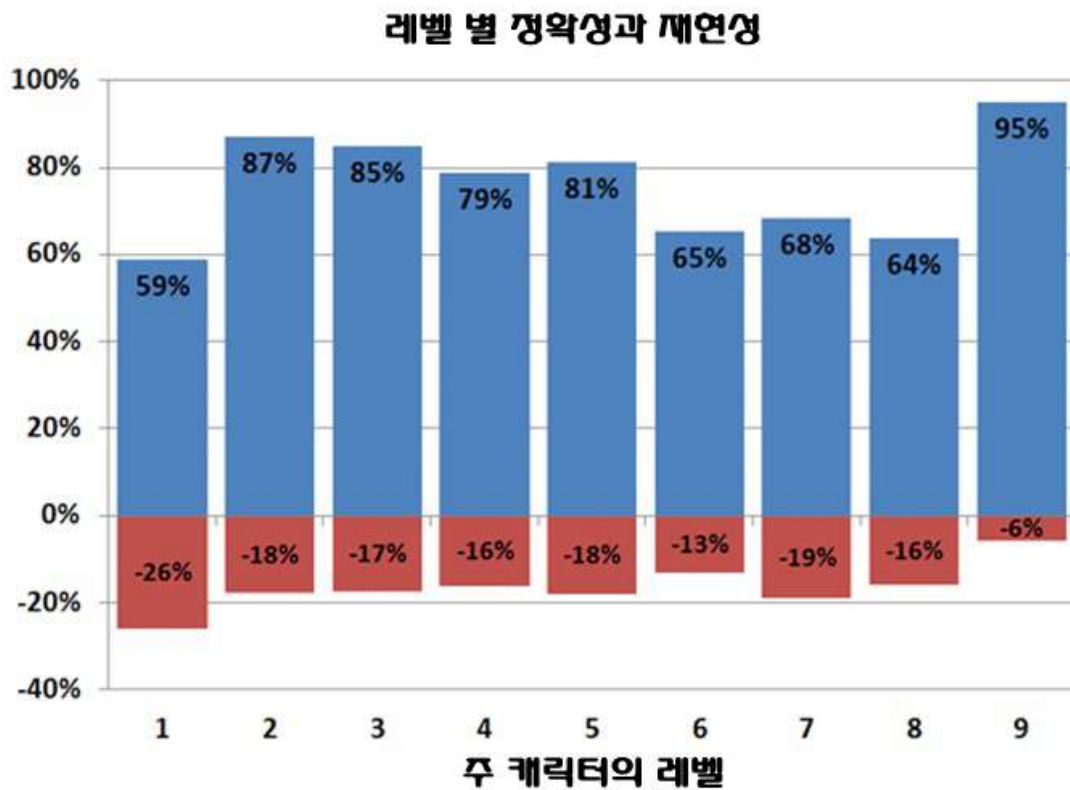
위의 예에서 우리는 자동 공격 자료를 수작업으로 분할했었다. 이 데이터를 파고 들기 시작할 때부터 SQL 서버의 자동화된 분할(discretization)은 미세조정 할 수 있고, 또 그렇게 해야만 한다는 것은 자명했다. 몇몇 버킷을 수작업으로 조정하자 의사결정 트리의 형상과 정확도가 큰 영향을 받았다. (다른 모델도 물론 큰 영향을 받았지만, 트리에서 가장 변화가 컸다)

나는 내 인생의 귀중한 **한 주일을 꼬박** 각각 30 이상의 차원을 가지는 9 가지의 데이터 마이닝 구조를 (각 레벨당 구조 하나씩, 총 9 개의 레벨) 미세조정 하는 데 보냈다. 버킷을 가지고 실험을 진행하자 몇몇 흥미로운 패턴이 드러났으며, 버킷 7 개와 8 개 사이의 차이는 정확도 면에 있어서 2 퍼센트에 이르는 차이를 보였다. 예를 들면, 사냥한 몹의

⁴ 참조링크: <http://www.amazon.com/Data-Mining-Microsoft-Server-2008/dp/0470277742>

수는 20 으로, 총 플레이시간은 12 시간, 레벨당 플레이시간은 7 시간 등으로 버킷이 안정화되었다.

미세조정을 통해서 긍정 오류를 상당부분 제거할 수 있었으며, 의사결정 트리를 신경망 정도 수준으로 강화할 수 있었다.



3 단계 결과: 마침내, 그럴싸한 모양새를 갖춘 결과를 얻었고, 플레이어들에 대한 많은 흥미로운 데이터를 모았다.

4 단계: 승리를 위하여!

솔직히 나는 정확도 면에서 최고점을 찍었다고 생각했었다. 새로운 측정법과 가설이 정확도를 높이는데 도움을 주지 않았고, 안정적인 모델이었다. 78 퍼센트의 정확도와 16 퍼센트의 긍정오류는 이탈 예측에 대한 작업을 충분히 시작할 만한 수치였다.

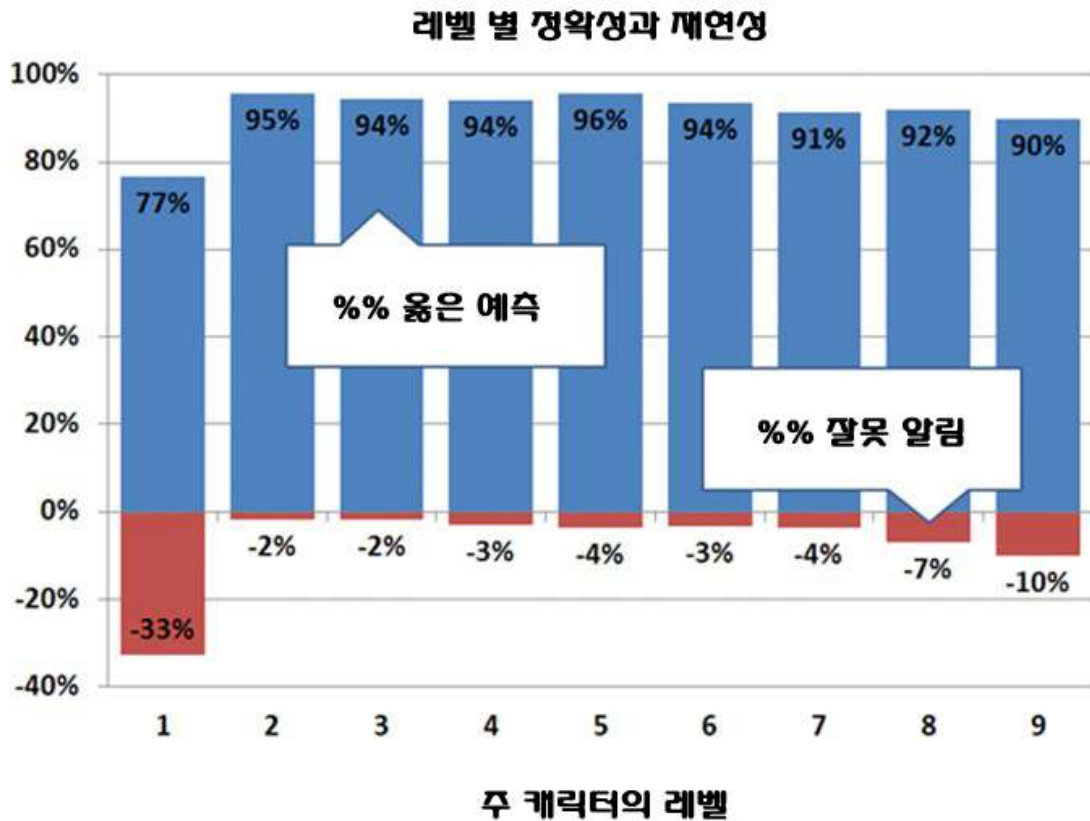
무료 가입이나 센 아이템으로 플레이어의 동기를 유발하는 것은 아마 효과가 없을 테지만(러시아에서 선물에 붙는 부가세와 비용을 고려해 볼 때), 플레이어들에게 이메일을 보내서 나쁠 것은 없지 않겠는가?

예상치 못한 선물: 석 달 간의 데이터 마이닝 작업을 통해, 우리는 패치를 몇 번 하면 옛날 데이터가 쓸모 없어질 수도 있다는 것을 깨달았다.

세 달 내내 새롭고 더 큰 데이터셋을 내려 받으며 나는 리프트 차트 상의 변화를 감지했다. 정확도와 재현율은 그대로였지만, 데이터는 미세하게 바뀌어 있었다.

ETL 프로시저를 처음부터 다시 작성하고, 3 개월간의 데이터셋을 굶주린 데이터 마이닝 괴물에 투입했다.

이 시점에 레벨 당 처리 시간은 1 분 미만이었고, 따라서 데이터셋이 커졌더라도 대기시간은 5 분이 늘어날 뿐이었다. 안타깝게도 수작업으로 이뤄진 미세 조정을 모두 다시 해야 했지만, 그림을 보자.



데이터셋 크기를 늘리자, 우리 모형의 효율에 엄청난 향상이 있었다!

레벨 1에서는, 도무지 제대로 된 결과를 낼 수가 없다. Avinash Kaushik 이라면 이렇게 말했을 것이다: “왔노라, 토했노라, 떠났노라.” 이런 플레이어는 캐릭터를 만들자마자 게임에서 이탈했기 때문에, 그 어떤 행위도 로그에 남아 있지 않았다.

위에 제시된 모든 숫자는 연대기적 데이터이며 우리 소중한 마이닝 모형의 학습 데이터셋이었다. 그러나 내가 무척 회의적인 사람인 탓에, 나는 실전 결과를 원했다! 그래서 우리는 오늘 갓 가입한 신규 유저를 예측 모델에 넣고, 결과를 저장해 보았다. 7 일 후, 우리는 일주일 전에 게임을 이탈할 것으로 예측된 플레이어와 그들의 실제 생활 양식을 비교해보았다. 그들은 실제 게임을 떠났을까, 아닐까?

행약도 결과	1	2	3	4	5	6	7	8	9
올바로 예측된 비이탈자	28	61	65	57	57	42	36	60	63
예측된 이탈자	424	569	682	571	522	365	360	246	132
부정 오류	308	11	26	15	21	33	32	19	11
긍정 오류	15	12	8	13	4	7	24	13	10
정확도 :	97%	98%	99%	98%	99%	98%	94%	95%	93%
재현율 :	42%	2%	4%	3%	4%	8%	8%	7%	8%

성과

원래 목표 - **게임에서 이탈할 플레이어 예측** - 달성은 성공적이었다. 높은 정확도와 재현율에 우리는 긍지와 자부심을 가지고 있다. 또한 결과물이 나온 시간도 적절했다는 점에 주목해야 한다. 오전 5:30 에 모델을 처리하고, 새 이탈자를 검출했으며, 우리가 아침에 출근했을 때 그들에게 보상을 줄 준비가 모두 돼 있었다.

우리는 두 번째 목표였던 **왜 플레이어가 이탈하는가?**에 대한 답을 알아내는데 성공했는가? 그렇지 않다. 그리고 이것에 나에게 가장 재미나는 부분이기도 하다 -- 플레이어가 이탈할 것은 매우 높은 정확도로 예측할 수 있지만, ~~왜~~ 그들이 이탈하는지에 대해서는 오리무중이다. 나는 왜 플레이어가 초반에 게임을 이탈하는지에 대한 가설을 나열하며 이 글을 시작했다:

- 종족과 직업.
- 플레이어가 이노바(Innova)의 다른 게임을 해본 적이 있는가? (우리엔겐 계정이 하나뿐이었다.)
- 얼마나 많은 캐릭터를 어떤 조합의 종족과 직업으로 만드는가?
- 체험기간 중 레벨당 사망 횟수와 총 사망 횟수
- 그 외 다수

우리는 60 여가지 이상의 개인별/게임별 측정 자료를 테스트했다. 그 중 어느 것도 이탈의 직접적인 원인이 되는 것이 아니었다. 그 어느 것도! 우리는 특효약을 발견할 수 없었다. 플레이어가 게임을 즐기지 못하도록 막는 그런 마법의 장벽 말이다.

연구에 있어서 핵심 측정 자료는 체험 첫날에 달성할 수 있었던 최고 레벨이라고 생각된다. 7 레벨 미만의 - 약 3 시간 가량의 플레이에 해당하는 - 플레이어는 이탈할 확률이 매우 높았다. 다음에 기술할 이탈 예측력이 높은 측정자료는 전체적인 활동에 대한 것이다:

- 레벨 당 몹 사냥 횟수
- 레벨 당 퀘스트 완료 횟수
- 레벨 당 플레이 시간
- 하루 당 플레이시간

결과

이 데이터 마이닝 프로젝트를 진행하는 데에는 3 개월의 시간과 두 권의 책과 어마어마한 열정이 필요했다. 우리 팀원 중 아무도 이쪽 방향의 일을 해 본 적이 없었던 것이다. 우리는 이노바의 튼튼하되 **수동적인** 분석 시스템 위에, 선구안을 지닌 **미래 예측 틀**을 만들어 냈다. 우리는 잠재적 이탈자 정보를 때맞춰 받으면 그들의 게임 플레이 경험을 향상시킬 수 있도록 고수준의 개인별 맞춤형 팁을 제공해줄 수 있다 (60 가지 이상의 측정자료가 모두 우리에게 데이터가 된다).

이 프로젝트는 <아이온>이라는 특정한 MMORPG 에 대해서만 이루어진 것이지만, 여러분이 보드시피, 다른 게임이나 웹 서비스에도 충분히 적용 가능한 일반적 측정 자료가 기여한 바가 크다.

이번 과제는 우리가 수행한 첫 데이터 마이닝 작업이었고, 2011 년 11 월에 끝났다. 그 후로 숙련된 플레이어의 이탈에 대해 우리의 현재 경험, 클러스터링과 세부화 분석, 플레이어들에 대한 깊은 이해 등을 기초로 이 자료를 처음부터 다시 작성했다. 그러므로, 데이터 마이닝이라는 모험은 계속 이어진다...