창직종합설계프로젝트1 [2] 팀별 report

머신러닝을 활용한 연주자 매칭   
웹서비스 개발

B611132 윤지원

B411040 김영실

**1.신기술 동향**

머신러닝은 다양한 분야에서 활용되고 있는데, 머신러닝의 활용 중 구현하고자 하는 서비스와 연관성이 높은 사용예시에 대하여 조사해보았다.

**Marketing Personalization**

The more you can understand about your customers, the better you can serve them, and the more you will sell.  That’s the foundation behind marketing personalization. Perhaps you’ve had the experience in which you visit an online store and look at a product but don’t buy it — and then see digital ads across the web for that *exact* product for days afterward. That kind of marketing personalization is just the tip of the iceberg. Companies can personalize which emails a customer receives, which direct mailings or coupons, which offers they see, which products show up as “recommended” and so on, all designed to lead the consumer more reliably towards a sale.

**Online Search**

Perhaps the most famous use of machine learning, Google and its competitors are constantly improving what the search engine understands. Every time you execute a search on Google, the program watches how you respond to the results. If you click the top result and stay on that web page, we can assume you got the information you were looking for and the search was a success.  If, on the other hand, you click to the second page of results, or type in a new search string without clicking any of the results, we can surmise that the search engine didn’t serve up the results you wanted — and the program can learn from that mistake to deliver a better result in the future.

**Sales Optimization**

Sales typically generates a lot of unstructured data that can ideally be used to train machine learning algorithms. This comes as good news to enterprises that have been saving consumer data for years, because it is also the place with the most potential for immediate financial impact from implementing machine learning. Enterprises eager to gain a competitive edge are applying ML to both marketing and sales challenges in order to accomplish strategic goals. Some popular marketing techniques that rely on machine learning models include intelligent content and ad placement or predictive lead scoring. By adopting machine learning in the enterprise, companies can rapidly evolve and personalize content to meet the ever changing needs of prospective customers. ML models are also being used for customer sentiment analysis, sales forecasting analysis, and customer churn predictions. With these solutions, sales managers are alerted in advance to specific deals or customers that are risk.

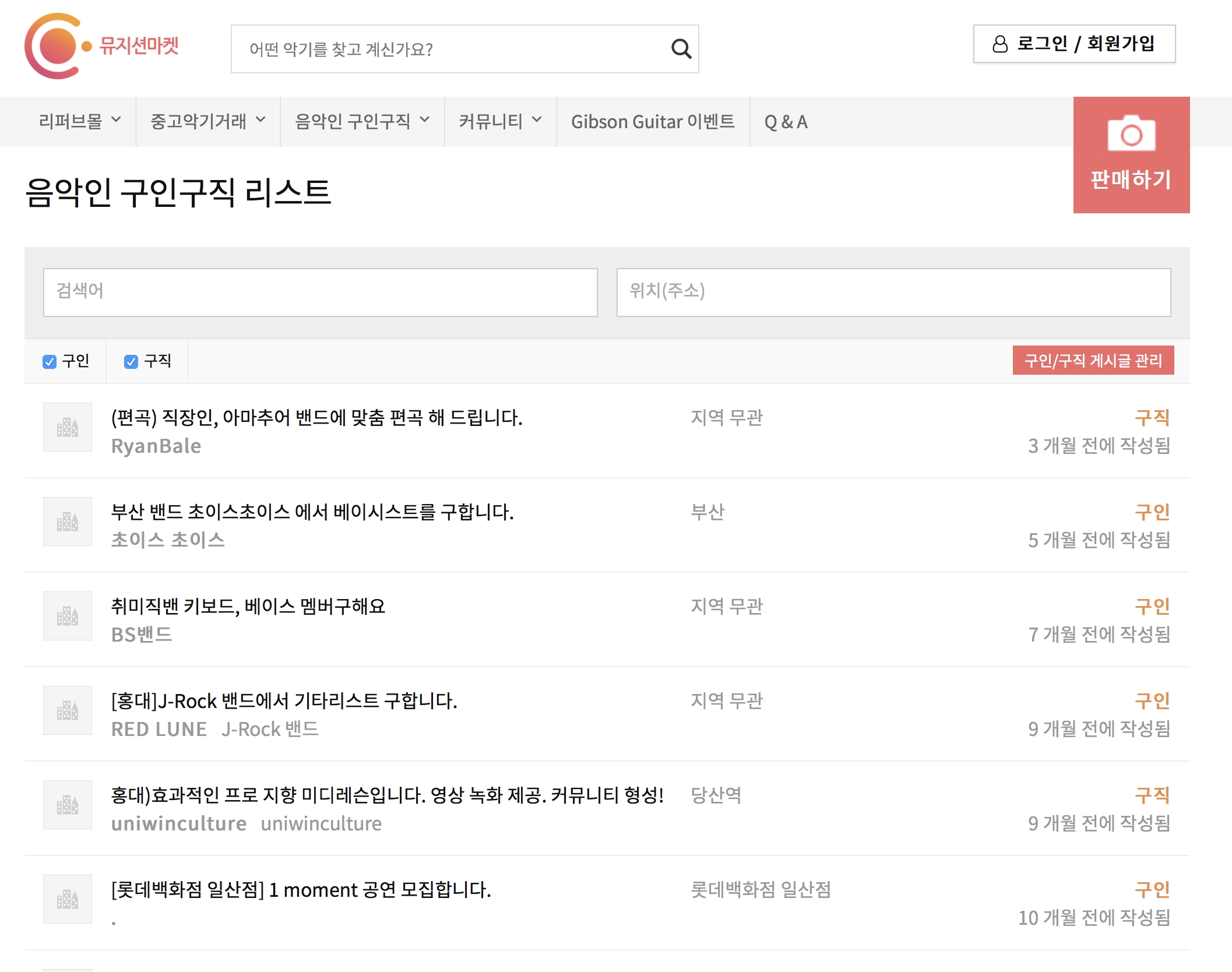
**2. 배경**

악기 연주자와 공연 주최자를 연결해주는 역할에 적합한 사이트가 없어 연주자를 구하는 사이트 만들게 되었다. 연주자가 직접 개별적인 모집공고를 확인하지 않아도 연주자에게 적합한 모집공고를 자동으로 추천하는 프로그램을 머신러닝을 활용하여 구현해보고자 한다.

**3. 기존 과제와 차별성**

기존의 음악인 커뮤니티 사이트는 음악인 구인구직 게시판을 활용하여 연주자를 구하고 있다. 사이트 중 뮤지션마켓과 뮬의 구인구직 예시를 살펴보자.

뮤지션마켓은 음악인 구인구직 리스트와 구인구직 등록으로 나뉘어 있다. 뮤지션마켓의 구인구직 리스트는 모집공고, 연주자의 지원공고, 연습실 대여공고가 분류되어 있지 않아 불편한 점이 있다.

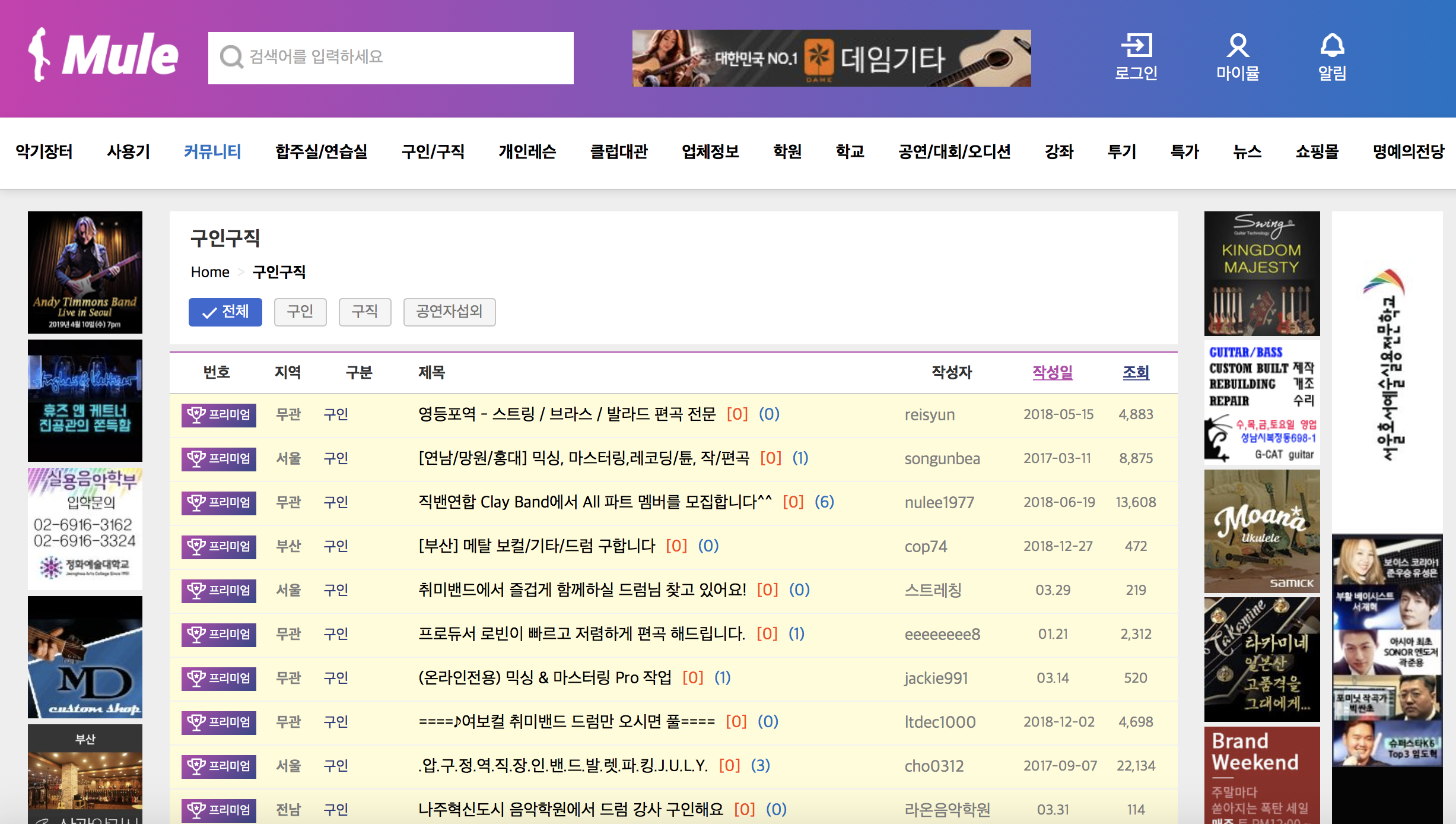


구인구직 리스트

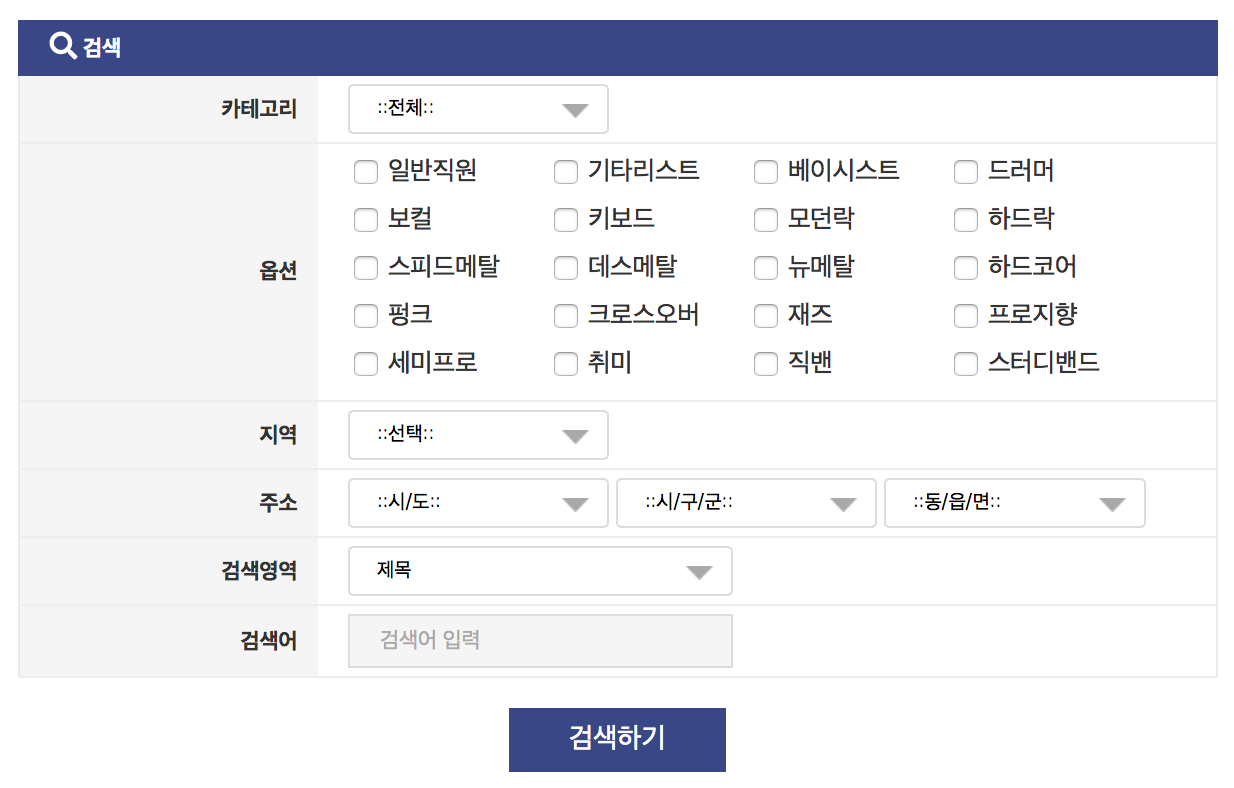
뮬의 경우에는 구인구직, 연습실 대여, 레슨모집 공고를 위한 게시판이 분리되어 있지만, 사용자들이 게시판을 분류해서 사용하지 않아 게시글이 혼재되어 이용이 불편하다. 각 게시판의 하단에 카테고리별로 분류하여 게시글을 검색할 수 있다.



구인구직 등록



뮬 구인구직 게시판



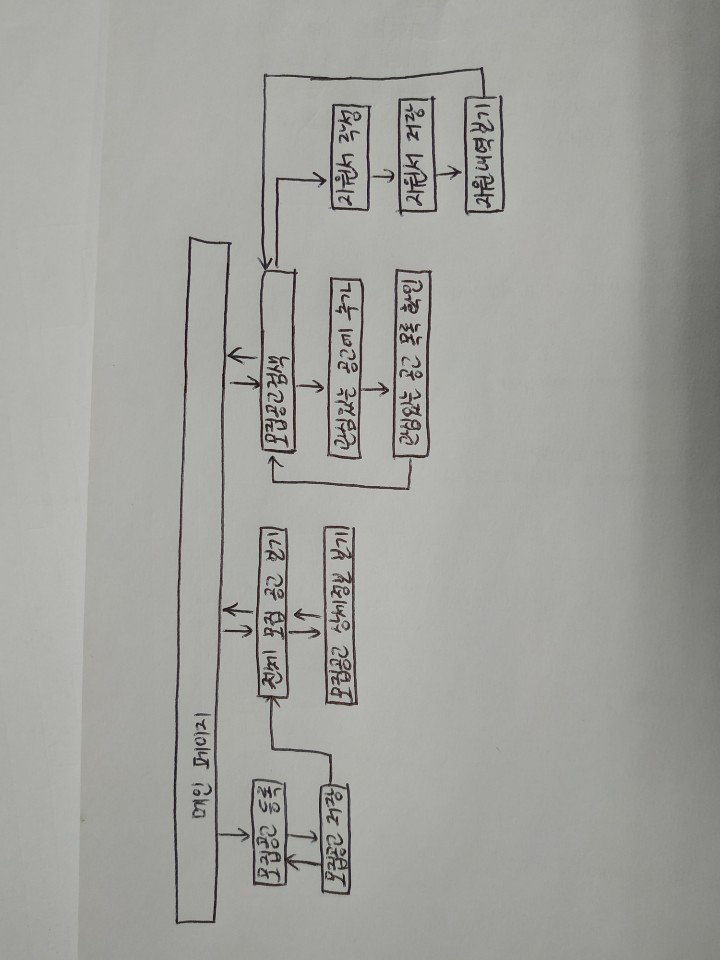
뮬 구인구직 게시판 내 검색

음악인 구인사이트에서는 모집공고 추천서비스를 제공하지 않는 경우가 많다. 따라서 사이트를 구현하고 추천알고리즘의 user-based filtering, item-based filtering을 활용하여 사용자에게 모집공고를 추천하려고 한다.

**4. 예상 문제점들 (risk분석)**

추천 알고리즘을 활용할 수 있을 만큼 많은 양의 데이터를 구하기 어렵다는 문제점이 있다. 사용자의 데이터는 구글 설문지를 이용하여 직접 데이터를 수집할 예정이며, 부족한 데이터는 수집된 사용자의 데이터를 조작하여 활용할 예정이다. 모집공고의 데이터는 외부에 유출하지 않는다는 전제로 다른 사이트의 정보를 연구목적으로만 활용할 계획이 있다.

**5. 시스템 구조**



**6. 개발환경**

개발에 사용할 IDE(통합개발 환경)은 Eclipse Oxygen을 사용.

Eclipse에 자바 언어를 사용하기 위해 java JDK8을 연동시켰고 Eclipse에서 바로 웹서버를 구동 시킬 수 있도록 tomcat 8버전 서버를 연동시켰다. 사용언어는 front end 웹서비스 구축 언어 html, css, javascript 가 있고 back end 웹서비스 구축 언어 java(JSP-=Java Server Page) 그 중에서 spring framework를 사용한다.

머신러닝에는 python의 TensorFlow나 R언어를 통해 협업 필터링을 구현.

DB에는 오라클을 사용.

Client는 윈도우를 사용, Server는 그래픽 요소가 적은 Linux를 사용.

**7. feedback**

<TensorFlow와 Soft Max>

**TensorFlow** : 구글(Google)에서 만든, 누구나 사용할 수 있는 딥러닝 오픈소스 라이브러리

TensorFlow API안의 하나의 함수로 SoftMax라는 함수가 있다.

**SoftMax** : SoftMax는 classification 알고리즘 중의 하나로, 들어온 값이 어떤 분류인지 구분해주는 알고리즘이다. 예를들어 A, B, C 3개의 결과로 분류해주는 SoftMax의 경우 결과값은 [0.7, 0,2, 0.1] 과 같이 각각 A, B, C일 확률을 리턴해준다. (결과값의 합은 1.0이 된다.)

Ex) TensorFlow tf;

tf.softmax(x);

<협업필터링>

유저기반 협업 필터링의 경우 유사한 사용자를 찾아내어 그 찾아낸 사용자의 구매내역이나 추천점수를 보고 아이템을 추천해주는 것이다. 우리의 경우에는 구매내역이나 추천내역 같은 것이 없으므로 유사한 사용자를 협업 필터링으로 찾아내어 그 사용자가 지원한 모집공고들을 추천해 줄 수 있다.

아이템 기반 필터링의 사용자가 어떤 공고모집을 눌렀을 때 그와 유사한 공고모집을 추천해주는 필터링이다. DB에서 같은 악기들끼리 모아서 그 모집공고들 중에서 유사도를 구해 제일 유사해 보이는 모집공고를 찾아내어 사용자에게 추천해 줄 수 있다.

소프트웨어 설계의 타당성을 고려해 본 결과 softmax함수를 사용하는 것은 적절하지 않다고 판단하여, 기존의 방식대로 협업필터링을 사용하기로 했다.

**8. 참고문헌**

<https://www.forbes.com/sites/bernardmarr/2016/09/30/what-are-the-top-10-use-cases-for-machine-learning-and-ai/#2c57e3b894c9>

<https://www.getfilecloud.com/blog/2018/08/top-5-use-cases-for-machine-learning-in-the-enterprise/#.XKIeFWYqHBI>