****

**5TS 집 가(집에 가!) 가구 매칭 보고서**

**개 발 자 : 김 선 종**

**개발 분야 : 가구 매칭**

**개발 기간 : 2018.10.01 ~ 2018.11.26**

**목차**

* **개발 동기**
* **개발 설계**
* **데이터 수집**
* **데이터 분석 가설**
* **데이터 가공 및 처리 내용**
* **머신러닝 활용**
* **개발 과정**

**개발 동기**

아직 인테리어를 하지 않은 상태거나 혹은 구매하고 싶은 가구와 어울리는 인테리어를 추천 받고 싶을 때, 가구를 업로드하면 인테리어를 보여주는 방식은 어떨까? 라는 생각으로 개발을 시작하였다.

**개발 설계**

가구 사진을 업로드 시, 사진에서 색을 추출하여 그 사진의 색들을 기준으로 인테리어를 보여주는 방식

**데이터 수집**

인테리어 이미지들을 수집한다.

크롤링 시 메모리의 과부화를 막기 위해 DB에 이미지를 저장함과 동시에 이미지의 이름과 이미지의 url 데이터를 저장하면서 크롤링을 한다.

크롤링 할 사이트가 무한 스크롤로 구현되어 있기 때문에 스크롤을 자동으로 내릴 수 있도록 구현한다.

**데이터 분석 가설**

인테리어 이미지에서 추출한 R, G, B(\*이하 feature로 정의) 값들 중 가장 많은 비율을 차지하는 feature 값과 업로드 된 이미지의 feature 값이 같은 범위에 있으면 어울리는 이미지일 것이다.

**데이터 가공**

이미지 데이터에서 R, G, B 값을 가장 많은 비율부터 5가지로 분류 추출한다.

5가지 중 가장 많은 비율의 R, G, B 값을 feature로 사용한다.

이미지에서 뽑아낸 feature 값들을 군집으로 분류한다.

k-means 군집으로 분류된 예측 값을 label로 사용하고 다시 SVM 머신러닝을 사용하여 지도학습을 시킨다.

업로드 될 이미지의 feature 값이 새로운 데이터이기 때문에 지도학습된 SVM 모델을 통하여 예측할 수 있도록 한다.

**머신러닝 활용**

그리드 서치를 활용하여 적절한 매개변수를 선정하고 모델로 활용

k-means 군집으로 분류된 예측 값을 label로 사용하고 다시 SVM 머신러닝을 사용하여 지도 학습을 시킨다.

PyCham에서 모델을 활용하기 위하여 모델을 저장한다.

업로드 될 이미지의 feature 값이 새로운 데이터이기 때문에 지도 학습된 SVM 모델을 통하여 예측할 수 있도록 한다.

업로드 되는 이미지 데이터를 모델로 분석하고 label을 받을 수 있도록 PyCham을 연결한다.

**개발 과정**

스프링 개발 환경을 이용하여 서버를 구축한다.

기능 1. 업로드 되는 이미지에서 feature를 추출하여 label을 분류한다.

분류 받은 label에 해당하는 이미지들을 업로드와 동시에 페이지에서 보여준다.

이미지 업로드 시 label을 자동 분류하기 위하여 SVM 머신러닝을 활용하였음.

기능 2. 가장 많은 비율을 차지하는 색을 우선으로 같은 label을 갖는 인테리어 이미지들을 먼저 보여준다.

기능 3. 가장 많은 비율 이외의 다른 색으로 더 볼 수 있도록 Ajax를 이용하여 클릭 시 보여줄 수 있도록 한다.

기능 4. ‘더보기’ 기능을 구현하여 페이징 처리를 했다.

기능 5. 로그인이 안되어 있으면 버튼의 클래스를 동적으로 만들어서 클릭을 하지 못하도록 하였음

기능 6. 이미지를 클릭하게 되면 해당 사이트로 이동하여 상세 정보들을 확인 할 수 있음