

데이터 분석 모델 기반 웹 서비스 개발자 채용 실기 과제

과제명: 시계열 데이터 이상 탐지 예측 모델을 연동한 간단한 웹 페이지 구현

○ 기능 구현 항목

- 사용자로부터 CSV 파일을 입력 받아 웹 페이지에 테이블 및 시계열 차트로 표현
- 시계열 이상 탐지 모델의 탐지 결과를 API 를 통해 전송 받아 웹 페이지에 차트/테이블로 표현
 - 예측 모델 결과를 전송하는 API 및 테스트 데이터는 함께 제공함 (아래 상세 설명 참조)

○ 화면 설명 및 상세 요구 사항

- 예상 결과 화면 (아래 번호 별 UI/동작 설명은 요구사항 참고)

데이터 입력

예측 결과

파일 입력

파일 선택

(입력 파일명 출력)

보기

입력

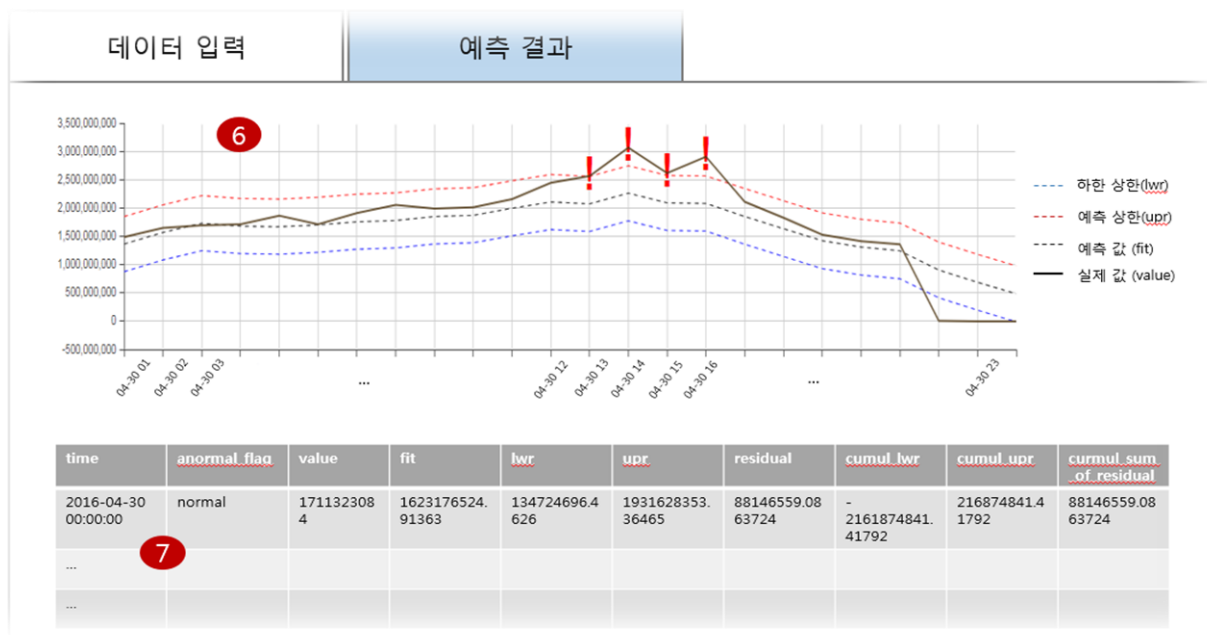
파일 입력 form

데이터 보기

업로드 csv 파일 출력 테이블

time	value	event1	event2
2016-04-30 00:00:00	1711323084	0	0
...			
...			

<그림.1>데이터 입력 탭 화면



<그림. 2> 예측 결과 탭 화면

▪ 상세 요구 사항

- ① 데이터 입력/예측 결과로 구분된 탭 패널로 기본 레이아웃 구현
- ② 데이터 입력 패널 상단에 파일 입력 form 이 위치 - 파일 선택 버튼을 눌러서 함께 첨부한 test_data.csv 파일이 업로드 되도록 기능 구현

- ③ '보기' 버튼 - 클릭하면 데이터 입력 패널 하단에 업로드한 csv 파일 내용을 (헤더/데이터 row 포함) 테이블(아래 4 번) 형태로 렌더링 한다.
- ④ 업로드한 csv 파일 내용 출력 테이블
- ⑤ '입력' 버튼 - 클릭하면 입력한 데이터 파일(test_data.csv)에 대한 이상 탐지 모델 결과를 API 서버에 요청 한다. API 서버 호출 및 응답 스펙은 아래와 같다.

- HTTP - POST 방식 (enctype="multipart/form-data")
- API URL: http://35.194.117.145:8080/upload_test_data
- 'test_data_file' 파라미터명으로 전달
- 입력 파일 형식: csv (test_data.csv 참고)

```
"time","value","event1","event2"
"2016-04-30 00:00:00",1711323084,"0","0"
"2016-04-30 01:00:00",1560383375,"0","0"
// 참고) 헤더, time: 시간, value: 해당 시간의 관측 값, event1, event2: 값에 영향을 주는 이벤트 변수
```

- API 결과 데이터 형식 예

```
[
  {
    "anormal_flag": "normal", // 데이터 이상 여부
    "cumul_lwr": -2161874841.41792,
    "cumul_upr": 2161874841.41792,
    "cumulative_sum_of_residual": 88146559.0863724,
    "fit": 1623176524.91363, // 예측치
    "lwr": 1314724696.4626, // 예측 하한
    "residual": 88146559.0863724,
    "time": "2016-04-30 00:00:00", // 시간
    "upr": 1931628353.36465, // 예측 상한
    "value": 1711323084 // 입력 실제값
  },
  ...
]
```

- API 호출 방식은 함께 첨부한 file_upload_test.html 파일 참고바람
 - 결과 응답이 완료 되면 alert 창 등으로 탐지 결과 수신 완료 Message 표기
- ⑥ 예측 결과 Line Chart - 예측 결과 탭 패널 상단에 입력한 테스트 데이터 파일로 요청하여 응답 받은 탐지 결과의 시간대 별 실제 값/예측 값/예측 상한/예측 하한 값의 Line Chart 를 출력한다. 실제 값이 예측 상한 값을 초과하는 경우 화면과 같이 "!"아이콘 등으로 이상 포인트가 강조 되도록 한다.
 - ⑦ 예측 결과 출력 테이블 - 예측 API 결과 JSON 리스트 데이터를 테이블 형태로 하단에 출력

○ 참고) API 서버에서 사용된 시계열 데이터 이상 탐지 예제 모델:

<https://github.com/ncsoft/timesuperin>

○ 기타 안내 사항

- 위 결과 화면은 예시이므로, 요구 기능에 대해 UI 디자인은 자유롭게 구현해도 무방함
- 각종 프레임워크/오픈 소스 라이브러리 사용 자유 (단 출처나 github 주소 명시할 것)
- 구현 코드에 대한 설명 필요 시 주석 작성
- 사용된 개발 환경/도구 명시 (런타임/개발언어/툴/프레임워크)
- 명시한 기본 요구 사항 외 추가로 구현하는 항목에 대하여 가산점 부여 (아래 내용 확인)

추가 구현 시 가산점 부여 항목

○ 시계열 이상 탐지 예측 모델 직접 구현

- 이 경우 API 호출 대신 본인이 구현한 탐지 모델을 호출하도록 기능 구현해야 함
- 개발 언어/머신러닝 프레임워크 사용 자유
- 데이터 셋: 첨부한 train_data.csv, test_data.csv 사용
 - 유사 형태로 직접 생성해서 사용해도 무방
- 모델링 과정 및 설명 문서 첨부

○ 모델 서비스 백 엔드 로직 및 HTTP Restful API 개발

- 언어/프레임워크/런타임 사용 자유
- (직접 예측 모델을 생성한 경우) 예측 모델 혹은 임의 예측 결과 데이터와 연동하는 API 직접 개발

○ HTML/컴포넌트 스타일링

- 관련 라이브러리 및 css 확장 언어 사용 가능
- UI 편의성/스타일 개선 및 다양한 효과 구현

○ JavaScript/CSS 프레임워크 사용 & 순수 JS/CSS 로 2 개 버전 구현

- 아래 2 개의 버전으로 구현이 가능한 경우
 - 기본 DOM 스크립트 API/CSS 만 사용하여 view 단 구현
 - 각종 모던 웹 프레임워크, JS/CSS 라이브러리 사용 구현