

「제11회 해양수산 비즈니스 · 공공서비스 공모전」
기 획 서(빅데이터 분석 부문)

1. 참가자 정보	
개인·팀·기업명	울지마(곽동현, 김도영, 김민정)
서비스(제품)명	울산항 선박 입출항 기록 및 차량 입출문 기록을 활용한 운송 차량 규모별 화물 적재 및 입출문 소요 시간 예측
공모 분야	<div> <input checked="" type="checkbox"/> V 해운·항만 물류 <input type="checkbox"/> 해양환경·기후·안전 </div> <div> <input type="checkbox"/> 해양·어촌관광·레저 <input type="checkbox"/> 낚시·귀여귀촌·양식 </div> <div> <input type="checkbox"/> 수산물 유통·안전·품질 </div>
서비스 유형	<div> <input checked="" type="checkbox"/> V 모바일(태블릿pc) 앱·웹 서비스 <input type="checkbox"/> 제품제작 </div> <div> <input checked="" type="checkbox"/> V 온라인 웹 서비스 <input type="checkbox"/> 영상 </div> <div> <input type="checkbox"/> 웨어러블디바이스 <input type="checkbox"/> 기타 () </div>

출처: 한국농수산식품유통공사(2015) 농수산물유통정보시스템

-

- ① 선박 및 차량의 데이터 병합을 통해, 선박 입출항 시간 사이의 화물차 데이터를 추출
- ② RandomForestRegressor를 사용하여 화물차 입출문 소요 시간을 예측
- ③ 레이블 인코딩을 통해 범주형 데이터를 수치형 데이터로 변환
- ④ 예측 모델은 선박의 척 수와 트럭 유형에 따른 입출문 시간을 비교적 정확하게 예측할 수 있음

[illegible]

- a) 울산항 선박입출항현황(210401-230331)

나스텍면 및 데이디 총회 세우치미문크정보센터() : ()

- ② 데이터 활용

- ③ 데이터 획득

실제 데이터 형태

[illegible]

사용 부두 목록(울산 본항)

순서	부두명	순서	부두명	순서	부두명	순서	부두명
1	1부두 01	13	6부두 04	25	SK5부두	37	일반부두 04
2	2부두 01	14	6부두 05	26	SK6부두	38	일반부두 05
3	2부두 02	15	7부두 01	27	SK7부두	39	일반부두 06
4	2부두 03	16	8부두 01	28	SK8부두	40	일반부두 07
5	3부두 01	17	8부두 02	29	SK부이 02	41	일반부두 08
6	3부두 02	18	9부두 01	30	SK부이 03	42	자동차부두 01
7	4부두 01	19	SK1부두 11	31	UTT부두	43	자동차부두 02
8	4부두 02	20	SK1부두 12	32	가스부두	44	자동차부두 03
9	5부두 01	21	SK2부두 01	33	남화부두	45	염포부두 01
10	6부두 01	22	SK2부두 02	34	일반부두 01	46	염포부두 02
11	6부두 02	23	SK3부두	35	일반부두 02	47	염포부두 03
12	6부두 03	24	SK4부두	36	일반부두 03		

선박 종류별 변수로 사용할 고유 번호 부여

번호	선박 종류	번호	선박 종류	번호	선박 종류
1	LNG 운반선	10	산물선(벌크선)	19	일반화물선
2	LPG 운반선	11	석유제품 운반선	20	자동차운반선
3	견인용예선	12	선박용도	21	철강재 운반선
4	관공선	13	세미(혼재)컨테이너선	22	케미칼 운반선
5	급유선	14	시멘트운반선	23	케미칼가스 운반선
6	기타 예선	15	신조선	24	폐기물 운반선
7	기타 유조선	16	압향 예선	25	폴컨테이너선
8	기타선	17	여객선		
9	모래운반선	18	원유운반선		

이 외에 입항일시와 출항일시가 비정상적인 데이터 중 상/하위 15% 삭제

b) 울산항만공사_최근 2개년 화물차 항만 출입 데이터

① 데이터 개요

- 시스템명 및 데이터 출처: 공공데이터 포털_울산항만공사
- 내용 : 21년 4월 1일부터 23년 4월 30일까지의 울산항 초소 입/출문 현황
- 획득 방법 : 공공데이터 포털 활용 신청을 통한 다운로드

② 데이터 활용

- 활용 빈도
- : 선박의 입/출항 일정과 화물차 입/출문의 관계를 학습시키는 데 사용
- : 모델 학습 시 사용 및 완성 이후 별도의 사용 용도는 없음
- 비중 : 화물차의 예측 체류시간(입문 후 출문까지의 소요 시간)
- 중요성 : 다양한 차종/차명의 차량이 입문 후 출문까지 걸리는 시간을 예측하기에 모델 학습에 필수적인 데이터

차량 무게별 크기 분류 (단위:ton)

크기 분류	해당 열 이름 : 내용
소	2.5톤 미만 해당 열의 차량 모두 1톤 미만
	1톤 미만
중	2.5톤 이상
	1톤 이상-5톤 미만
대	5톤 이상
	컨테이너 차량 특수차량(중장비)

③ 데이터 획득

- 지속성 : 23년 3월 31일 기준 2개년 데이터로 한정
- 활용범위 : 화물차와 선박의 관계를 예측하는 모델 학습시키는 데 사용
- 가공 가능성
- : 차량 입출문 가능 초소명 중 울산 본항만 사용
- : 차종의 개수를 줄여서 1~5톤 등 범위로 줄이고, 대-중-소 분류 열을 만들어 차량을 분류
- : 차량에 오타 및 똑같은 차량인데 다르게 분류된 부분 정리
- : 출문시각에서 입문시각을 감산하여 화물차의 체류시간을 계산하여 새로운 열로 추가
- : 데이터 전처리를 통한 예측 정확도 향상 가능 (입/출문 시간 누락 데이터 삭제, 화물차량 규모 라벨링 등)
- : 입문시각, 출문시각 열의 비어있는 행 삭제
- : 실제 하역 과정 절차의 시간이 최소 20-30분 소요이기에 예측 정확도/신뢰도 향상을 위해 입문과 출문 시각의 차이가 너무 작은 하위 15% 행 삭제

c) 해양수산부_선박운항정보

① 데이터 개요

- 시스템명 및 데이터 출처: 공공데이터 포털_해양수산부
- 내용 : 조화 시작/종료 날짜, 항구를 입력했을 때 선박의 일정 공지
- 획득 방법 : 공공데이터 포털 활용 신청을 통한 인증키 발급

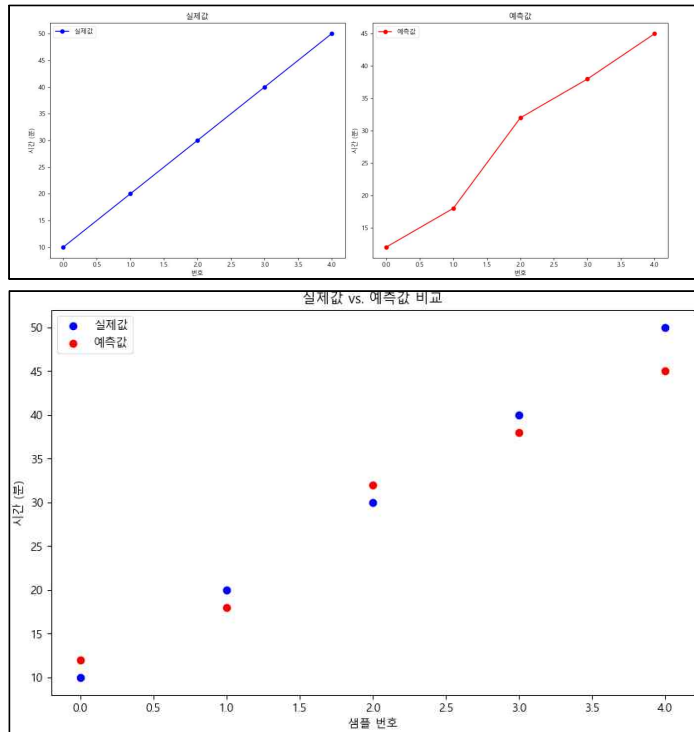
② 데이터 활용

- 활용 빈도 : 예측 모델에서 화물차의 입항 월과 일을 입력받을 때마다 해당 일자에 입항하는 선박이 몇 척인지 카운트
- 비중 : 입항 선박의 수가 증가할수록 항구에 출입하는 차량의 대수 역시 비례하여 증가하기에, 예측 과정에서 선박의 척 수는 필수 요소임.
- 중요성 : 실시간 조화를 통해 실시간으로 예측하기 위한 필수 데이터
- 데이터 변수명(feature값) 및 예측값(target명)

변수명	입문시각_연도	입문시각_월	입문시각_일	입문시각_시간	차종	선박_갯수
target명	걸린시간					

- 사용 모델 : Scikit-learn의 라이브러리 RandomForestRegressor를 사용한 모델 학습 및 평가

실제값과 예측값 그래프 비교 및 차이



③ 데이터 획득

- 지속성 : 조회 시작일자과 종료일자를 입력하여, 스케줄이 입력된 미래의 선박 입출항 현황까지 조회 가능한 api이므로 매우 지속적
- 활용범위 : n월/n일 차량이 입문한다고 입력값이 들어오면, 해당 날짜의 울산항에 들어올 예정인 선박의 수를 카운트해서 예측 모델에 입력.
- 가공 가능성 : 울산항으로 검색했을 때 울산본항, 울산신항, 온산항, 미포항 등 울산의 모든 항구가 포함됨. 이에 <laidupFctyNm>을 사용하여 부두를 구별하고, 울산 본항에 해당하는 부두에 입/출항하는 선박만 카운트 (제1부두, sk1부두, 그 외 기타 등)

1.3) 기존 서비스와의 차별성 및 독창성

(1) 기존 서비스와의 차별성

a) 울산항만공사 PortWise

- ① 항만 모니터링 등 입항 선박 관련 데이터에 서비스가 집중되어 있기에, 화물차량의 운송 계획 경향에 대한 직접적인 예측 자료로 사용될 수 없음
- ② 항만 데이터 플랫폼으로 활용되고 있기에 해운 물류 기업은 서비스의 직접적인 수혜 대상이 될 수 있으나, 화물차량 운전 기사에게는 서비스의 혜택이 간접적임

b) 부산항만공사 차량 반출입예약시스템(VBS, Vehicle Booking System)

- ① 트럭 운전자가 사전에 반출입할 화물과 시간을 예약하는 시스템으로써 예약을 바탕으로 터미널의 차량 집중 시간을 예측할 수 있으나 예약이 이루어지기 전에는 특정 시점에서의 경향성은 예측이 불가함
- ② 대기하는 화물차량의 규모(소형, 중형, 대형 등)에 따라 화물 적재 시간이 차이가 나는 점을 반영하지 못하고 있으며 이는 불필요한 교통 체증으로 이어짐

(2) 분석된 정보의 창의성

a) 기술

① 데이터 전처리

- 울산항 중 울산 본항으로 차량 출입 초소명과 선박 입출항 부두 수를 제한
- 차종 및 차명 재라벨링 후 이에 기반하여 대-중-소 열을 추가로 만들어 차량을 분류
- 누락 데이터의 열 개수 확인 후 삭제 및 상/하위 15%인 차량 입/출문 시각 값 삭제

② 학습

- Scikit-learn의 라이브러리 RandomForestRegressor를 사용한 모델 학습 및 평가

③ 출력

- Joliblib을 이용한 pkl 파일 형식에서의 모델 저장 및 모델 로드 후 결과값 출력
- flask 서버 구현을 통한 프론트엔드와의 연결 및 결과값을 직관적으로 볼 수 있도록 시각화

b) 서비스

- ① 기존의 서비스는 선박의 운영 시간에 집중되어 실질적으로 항구 이용객 중 차량 소지자들에게 도움이 되지 않았던 서비스. 본 공모안의 서비스는 기존의 선박운항 자료와 이전에는 고려되지 않았던 차종별 차량 입/출문 자료를 기반하기 때문에, 서비스 대상 자체를 해운 관계자를 넘어 항구 입출문하는 모든 차량 이용객으로 확장.
- ② 웹 환경 및 모바일 환경에서의 직관적인 화면 디자인을 통하여 서비스 제공 계획.
- ③ 실제 서비스 화면
- 웹 서비스 화면



예측 기능 서비스 화면



- 모바일 서비스 화면

메인 화면	회원가입/로그인 화면	인근 주유소 정보 제공 화면
질의응답 게시판 화면	예측 기능 서비스 화면 1	예측 기능 서비스 화면 2

2. 사업화

2.1) 분석된 정보를 활용한 사업계획

(1) 활용방안 및 파급효과

a) 활용방안

: 울산항 선박 입출항 기록 및 차량 입출문 기록을 활용한 화물차량 규모에 따른 특정 시점의 화물 적재 및 입출문에 소요되는 시간의 예측 정보를 제공하는 반응형 웹 서비스 개발
(모바일 및 PC 서비스 개발)

b) 파급효과

- ① 화물차량 운전자 관점 : 불필요한 대기 시간을 줄이고 효과적인 운송 계획 수립 가능
- ② 주변 교통 관점 : 울산항 입문을 위해 대기하는 화물차량 수를 줄여 주변 교통 체증 개선에 기여

(2) 창업 추진 의지 및 창업 계획

a) 활용방안

- ① 화물 차주 대상 무료 배포 애플리케이션 제작
- ② 애플리케이션 내 광고 게시를 통하여 수익 창출

2.2) 자유 타이틀 기재

※ 출품작에 대한 기타 추가 내용이 있을 경우 작성

3. 기타

3.1) 추가로 제공이 필요한 데이터(필요 시 작성)

- 기존에 제공된 데이터 외에 추가로 제공이 필요한 데이터명(목록)

※ 기획서 작성 시 유의사항

- 이미지 파일은 문서 내 포함하되, 분량은 10페이지 이내로 작성
- 하단에 제시한 목차 외 추가 내용이 있을 경우 별도 타이틀을 기재하여 작성
- 심사표를 참조하여 해당 내용이 누락되지 않도록 구체적으로 작성

* 심사위원 종합평가 의견에 따라 수상작이 적합하지 않을 경우 시상규모를 축소 또는 변경 가능