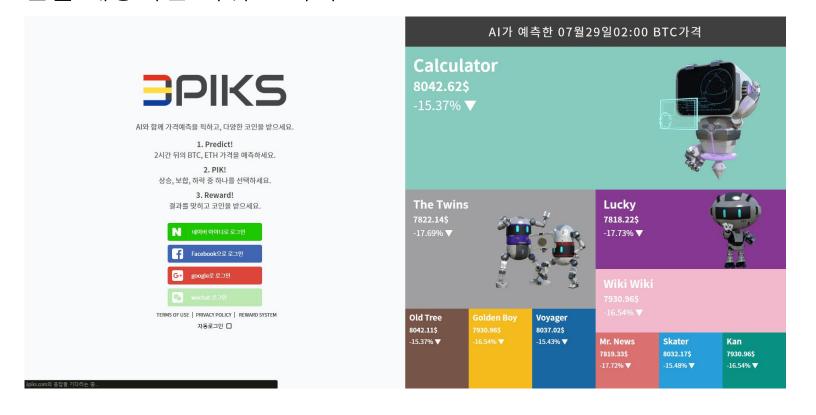
# 3PIKS

- ❖ 프로젝트 진행일 : 2018.10.23 ~ 2019.02.08
- ❖ 소속 회사명 : (주)뉴럴비씨

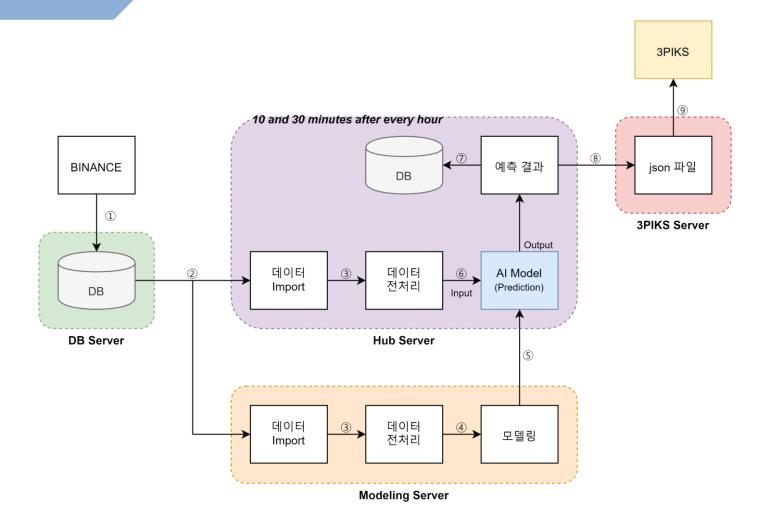
- ❖ 개발 환경 : Naver Cloud Platform Server (8 vCPUs, 60GiB Memory, 2 P40 GPUs, Ubuntu)
- ❖ 주사용 기술: Machine Learning, Python, MySQL, MongoDB

### 프로젝트 개요

❖ 유저들이 Al Model의 예측 결과를 참고하여 암호화폐의 가격 변화를 예측하고, 이에 대한 보 상으로 Al코인을 제공하는 리워드 서비스



### 파이프라이



### 파이프라이

#### 요약

- ① DB에 암호화폐 가격 정보 Insert
- ② DB에서 암호화폐 가격 데이터 Import
- ③ 데이터 전처리 진행 (중복 데이터 처리, Feature 추가 등)
- ④ 전처리된 데이터로 모델링 및 모델 저장
- ⑤ 저장된 모델 Hub Server로 전송
- ⑥ 매시 10분, 30분에 예측 시행 (한국기준 홀수시간 : BTC, 짝수시간 : ETH)
- ⑦ 예측 결과 DB에 저장
- ⑧ 예측 결과를 json 파일로 저장하여 3PIKS Server로 전송
- ⑨ 3PIKS에 예측 결과 반영

## 주요업무

#### Modeling

❖ 목표 : 2시간 후 암호화폐 가격 변화 예측 (Accuracy 높이기)

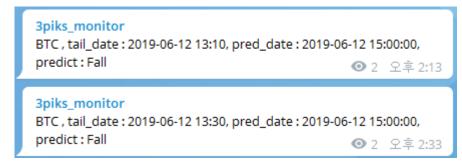
❖ Input : 암호화폐 가격 정보 Output : 상승, 보합 or 하락

- Modeling
  - ❖ ARIMA 모형과 Logistic Regression 모형 등을 사용하여 대략적인 성능 체크 및 Feature 연구
  - ❖ Boosting 계열의 알고리즘으로 모델링
    - ❖ 시계열 데이터에 많이 사용되는 RNN, LSTM 모델을 사용해보았지만, 성능 개선이 되지 않아 Bias를 줄이고 Feature를 연구할 수 있는 Boosting 모델 사용
    - ❖ Feature Importance를 참고하여 Feature 연구
  - ❖ 성능 개선을 위해 다양한 모델을 생성하여 Voting

## 주요 어무

#### 프로세스 관리 및 유지보수

- ❖ 파이프라인의 ②~® 프로세스 관리 및 모니터링
  - ❖ 적절한 시간에 프로세스가 작동하도록 Crontab 이용
  - ❖ Telegram을 통해 적절한 시간에 프로세스가 작동하는지 확인



- ❖ 다수의 Model 관리
- ❖ Error 발생시 원인을 파악하기 위해 log 파일 관리

## 주요어무

### 주요 이슈 및 해결

| 이슈                              | 해결                              |
|---------------------------------|---------------------------------|
| 클래스 불균형 문제(보합 label이 거의 없음)로 보합 | Accuracy 높이는 것이 목표이기 때문에 보합을 거의 |
| 을 거의 예측하지 않음                    | 예측하지 않더라도, 상승과 하락을 더 잘 맞출 수 있   |
| ⇨ 재샘플링 방법으로 클래스를 균등하게 하여 모델     | 도록 모델링                          |
| 링했지만, 오히려 Accuracy 감소           |                                 |
| Overfitting                     | 적절한 Parameter 값 지정 및 다양한 모델 앙상블 |

### 선과

❖ 유저들에게 20개의 AI Model 중 가장 높은 적중률로 암호화폐 가격 변화 예측 서비스 제공



