

# 기업신용평가 측면에서 빅데이터 분석 및 활용 방안

강응찬

---

## 목차

- 기업 신용평가에서 빅데이터 적용요소
- 빅데이터 적용 단계
- 빅데이터 수집 및 처리
- 빅데이터 분석 모델 구현
- 분석 모델 평가
- 기업 신용평가에 결과 반영

## 기업 신용평가 요소와 빅데이터 적용 요소

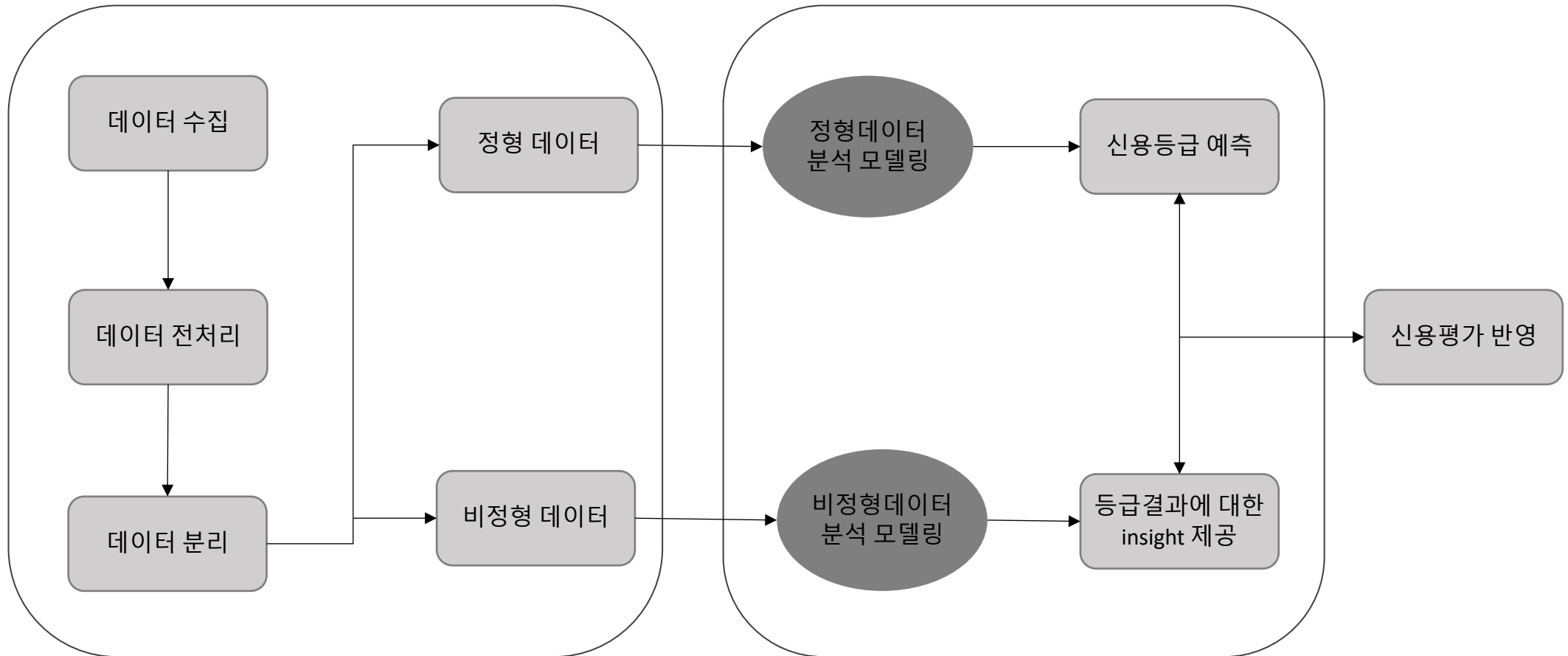
재무 모형	수익성
	안전성
	부채상환능력
	유동성
	활동성
	성장성
비재무 모형	산업위험
	경영위험
	영업위험
	재무위험
	신뢰도

빅데이터 분석 모델 적용

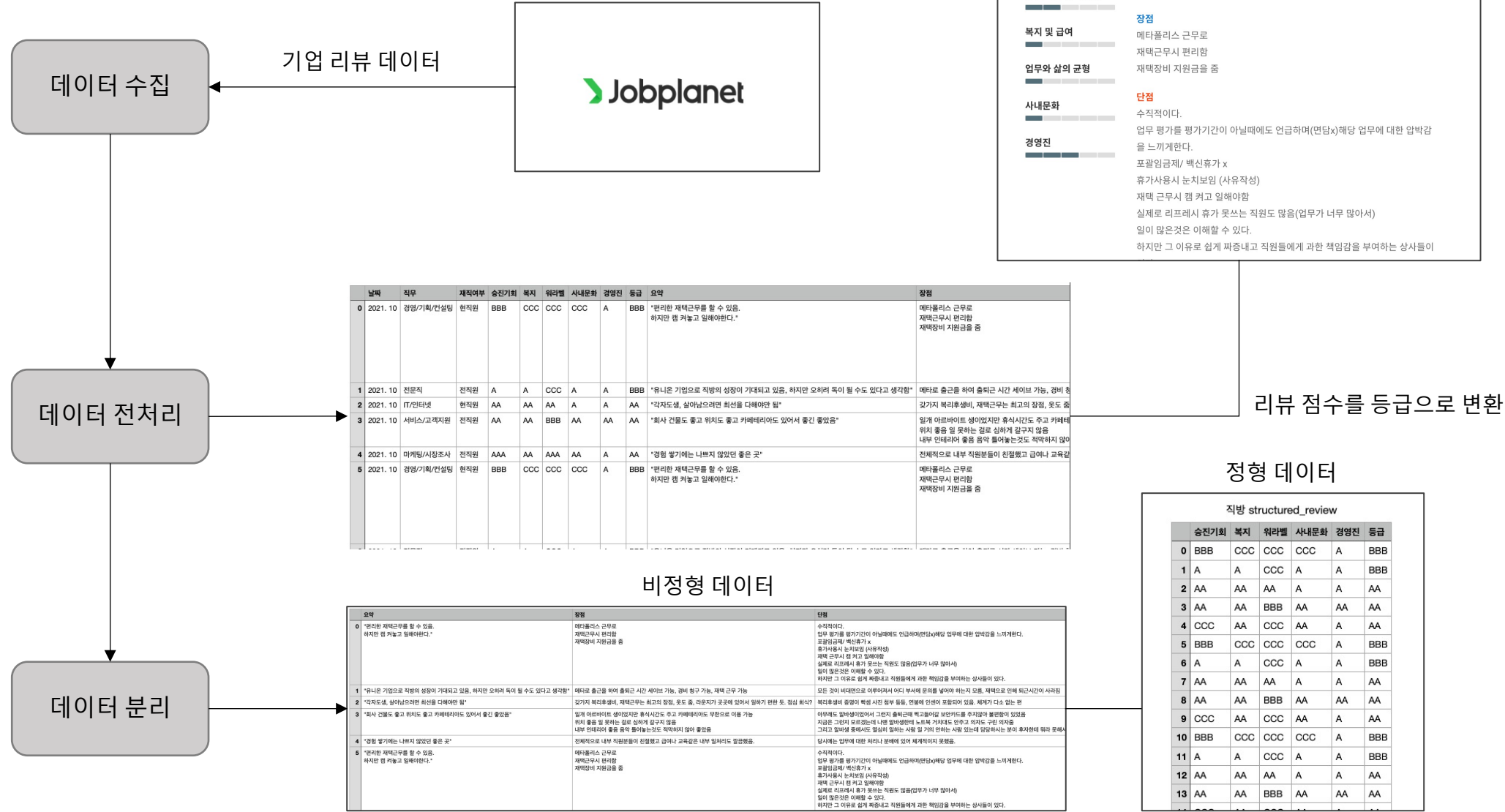
경영관리 능력 및 경영 안전성 평가를 위한 빅데이터 분석 모델을  
제공하여 비재무 모형의 평가 정확성과 신뢰 확보

💡 keyPoint: 기업 리뷰 데이터를 대량으로 수집하여 평가 요소별 관계에 따른 기업 신용등급 예측

## Big Data Modeling Process



# 데이터 수집 및 처리



# 정형 데이터 분석 모델 구현

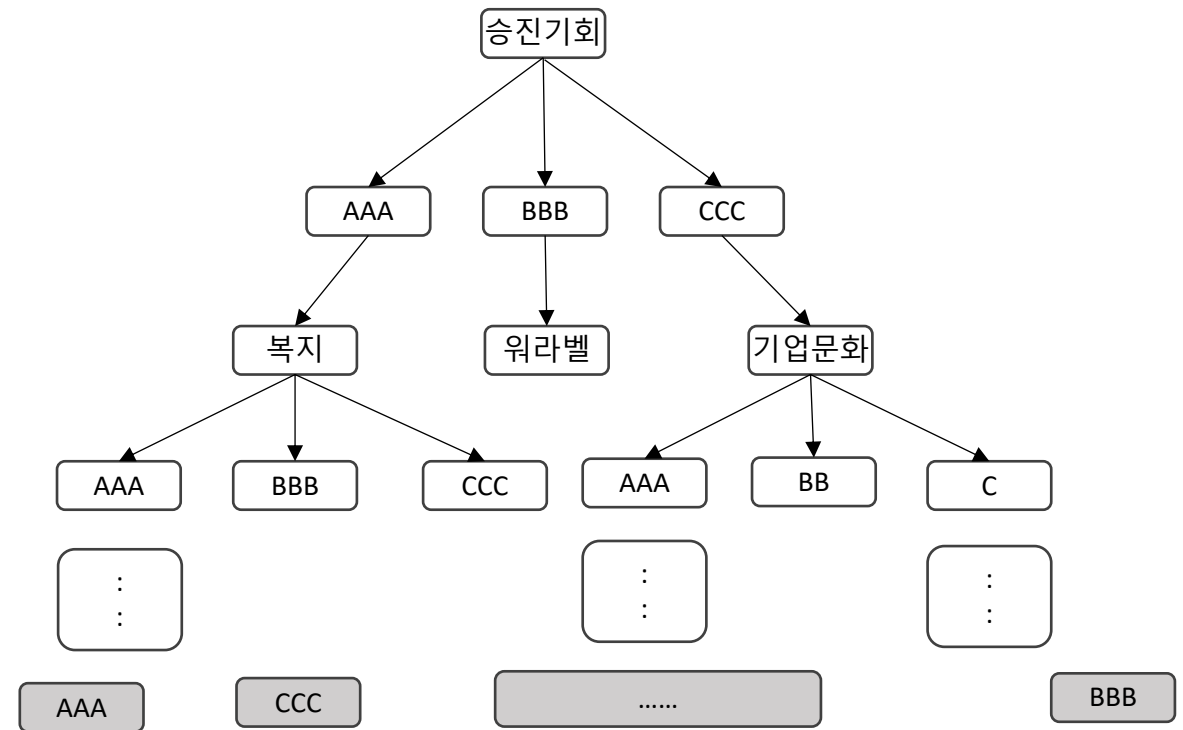
## 💡 목적

기업의 리뷰를 데이터로 활용하는 데서 가장 큰 문제는 신뢰성이다.

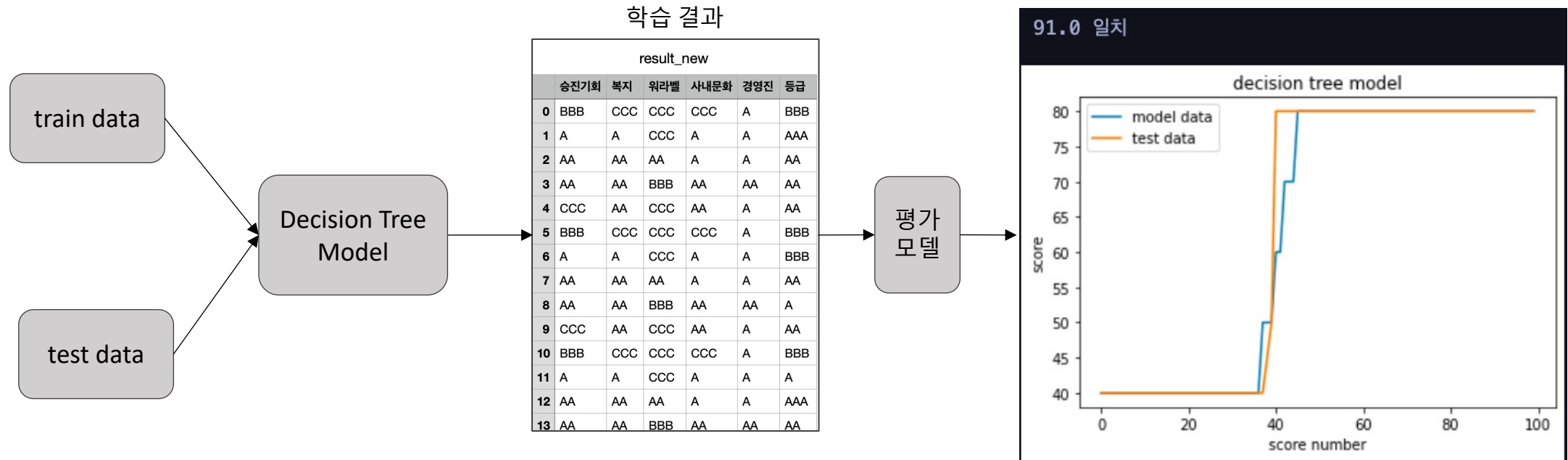
Decision Tree Algorithm을 이용하여 기업리뷰에 대한 신뢰도를 검사

### • 모델설명

- 기업의 리뷰 데이터중 90%를 train data로 이용하여 decision tree model 구현
    - 모델은 다음과 같은 것을 예측할 수 있다.
    - 승진기회 AA, 복지 BB, 워라벨 A, 기업문화 AA 이면 → **최종 등급은 AA**
  - 나머지 10% 데이터를 test data로 정하고 최종등급 값을 따로 저장
  - 최종등급이 없는 test data를 train 데이터로 만든 모델로 테스트하여 최종등급 생성
  - 따로 저장해 두었던 test data의 최종 등급과 새로 만들어진 등급이 90% 이상 일치하면 기업의 리뷰데이터의 신뢰를 입증
- 각 기업의 리뷰는 최종 점수를 먼저 주고 각 항목별로 따로 점수를 주는 방법으로 작성된다. 만일 리뷰데이터가 신뢰를 보장한다면 각 항목별 점수와 최종 점수는 반드시 관계가 있어야 한다는 아이디어에서 만들어진 모델.



## 모델 평가



x축: 평가 점수의 순서

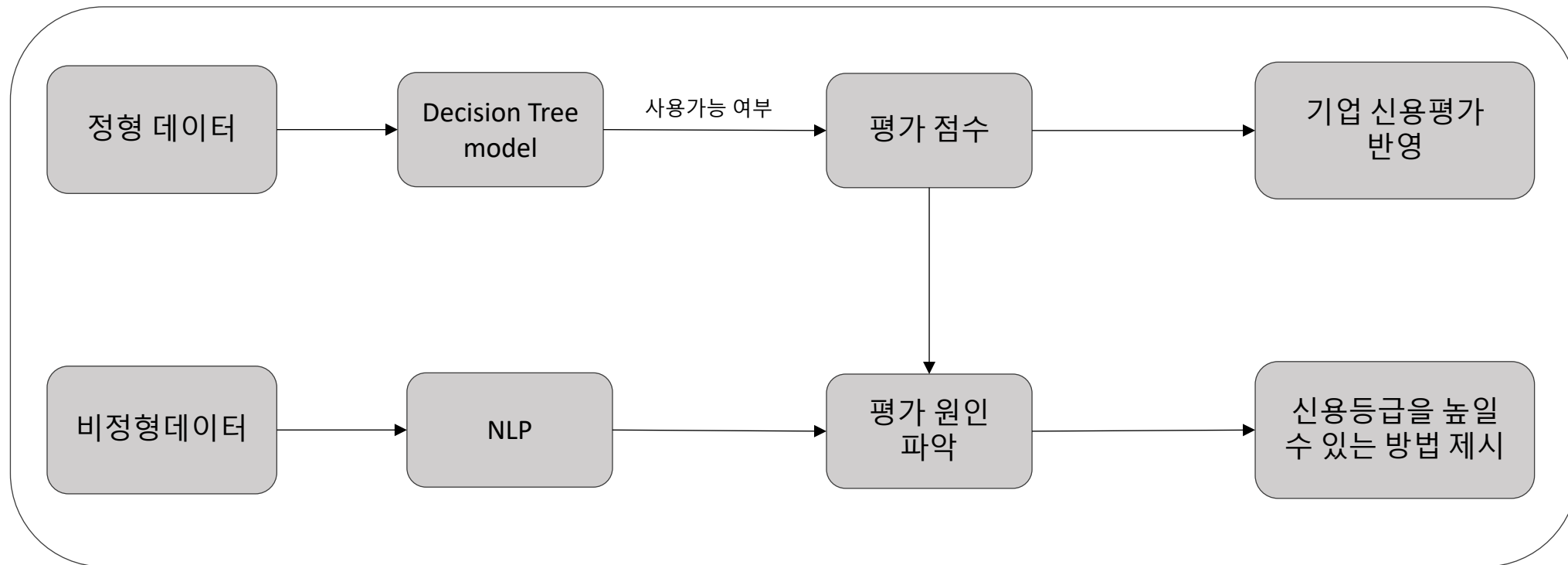
y축: 최종 리뷰 점수

빨간색 : 테스트 데이터

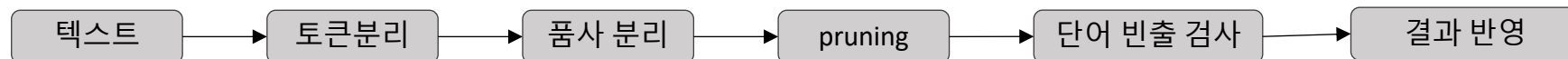
파란색: 모델에서 생성된 데이터

그래프를 보면 두 그래프가 거의 겹친 모습을 볼 수 있다.(91%일치)

## 신용 평가에 반영



### NLP



NLP는 “감성분석”이나 “텍스트 요약”과 같은 복잡한 알고리즘을 사용할 수도 있지만 위와 같은 단순한 알고리즘으로 구현할 수도 있다. 요약하면 텍스트를 품사로 분리하고 필요한 품사(예: 명사, 동사, 형용사 등)만 추출하여 그중 가장 많이 등장하는 단어로 정렬시킨다. 해당 기업의 평균 리뷰가 C 등급이고 NLP 결과에서 “수직구조”, “휴가” 등의 단어가 가장 많이 등장한다면 해당 기업이 점수가 낮은 원인을 쉽게 파악할 수 있다.



---

**감사합니다**