1. **과제(꼭 하세요!!!!!)**
2. **팀명: 오니기리와 여규동**
3. **프로젝트 목표**

* 교수님 피드백

1. **진행 사항**
2. 교수님 피드백
   1. 제목 관련
      1. 제목

머신러닝을 활용한 한계기업 중 회생기업의 재무적 특성 비교를 통한 정책적 제언

— 경제 충격기간과 정상기간의 비교를 중심으로 —

예시: 딥러닝을 활용한 ~ 이 많기 때문에 이런 식으로 제목을 만들자.

교수님이 주신 경기 주기별 분류 기준을 활용하는 것도 좋을 것 같다.

정상기간 비정상기간, 충격기간, 정상기간 등 기간 명칭을 정의하자.

머신러닝을 활용한 경제충격 기간 별 한계기업의 회생 주요 피처에 대한 정책적 제언

머신러닝을 활용한 경제 충격에 따른 한

기간 정의 > 관련 논문 활용하면 좋을 듯 합니다. 자료 찾아서 적용하자. 기간 분류에 대한 좋은 기준이 있다면 그것을 활용하자.

정상기간을 전체를 가져올지 기간을 맞춰 일부만 가져올지도 이야기 해야함.

한계기업에 대한 정의와 기간에 대한 정의 > 우리가 찾은 것이 있으니 이것을 활용하자.

* + 1. 기간 설정의 객관화 -> 자료를 통해서

Ex: 경기 침체와 확장의 구분: 통계 자료 기반 결정 등

* + 1. 충격이 기업 규모에 따라 차이가 날 것으로 보이기에 시기에 따른 차이도 보여주면 좋을 것 같다.
  1. 추가 피드백
     1. 외부적 충격이 있을 때의 기업 부실에 영향을 주는 변수를 선정하여 코로나 때 어려워진 기업들 중에서 앞으로 정상상태로 돌아갈지, 부도가 날지 예측한다.
     2. 기간은 2가지로 분류하여 분석하는 것이 좋을 것 같다. -> 정상, 충격 기간 분류
     3. 1시기, 2시기 피처값이 따로 있으니
     4. 특수기간 피처 선정, 정상상태 기간 피처 선정 2가지로 분류하여 데이터셋을 나누면 좋을 것 같다. 그래서 선정 피처의 차이를 보여주는 것도 좋을 것 같다.
     5. 데이터 셋에 따라 기업의 부도에 대해서 시장의 충격이 왔을 때와 아닐때와 차이가 있기 때문에 시기에 따른 기업 평가를 다르게 해야 한다는 정책적 제언에 활용이 가능할 것을 보인다.
     6. 판데믹 기간에 해당 결과를 제언한다는 방식으로 현재에 적용할 필요성을 제언한다.
  2. 한계기업 정의 관련
     1. 굳이 한국은행 정의 보다는 기타 정의에 레퍼런스를 해서 자체 한계기업을 정의할 수 있을 것 같다.
     2. 이상호 자본시장 연구원 자료: 우리나라 실정에 맞는 새로운 한계기업 정의가 필요할 것으로 생각된다.
     3. 만성적인 한계기업에 대한 정의도 하면 좋을 것 같다.

BOK 연구 보고서에서 만성적 한계기업의 정의를 관측 기간 내에서 1번이라도 한계기업으로 분류된 기업으로 정의한다.

* + 1. 한계기업을 선정한 이후 여기서 별도로 만성적 한계기업을 따로 뽑는다.(전체 데이터 세트에서든 기존에 정한 한계기업 데이터 세트에서든.)
  1. 기업 분류에 관한 정의
     1. 중소기업의 경우 데이터의 신뢰성 문제가 발생할 것으로 생각된다.
     2. 상장 기업을 활용할 경우 기존 연구와의 차별성이 힘들 것 같다.
     3. 스타트업의 경우 관련 재무 정보를 벤처협회에서 구할 수는 있을 것 같다.

1. 한계 기업 정의
   1. 한국은행:
      1. ‘3년 연속 이자보상배율(영업이익/이자비용) 1 미만’
   2. 블룸버그:
      1. 레버리지 비율(부채/자본) 100% 이상
      2. ROE 5% 이하인 기업
   3. 중진공
      1. 2년 연속 적자기업(영업손실) 중 자기자본 전액 잠식 기업(= 완전잠식)
      2. 2년 연속 매출액이 50% 이상 감소한 기업
   4. 신용보증기금:
      1. 최근 2개년 연속 당기순손실 실현
   5. 기업은행
      1. 당기 차입금의존도 70%(차입금/총자본) 이상

합집합으로 ‘오니기리 한계기업’ 정의

1. 조사 기간 설정
   1. 코로나 시기를 생각하면 2020~2021 데이터 포함 필수적
   2. 타겟: 코로나 시기에 한계기업으로 분류된 애들을 볼 것인지, 17~19년도를 기준으로 해도 되고…
   3. 만약 코로나 시기를 고려한다면 과거 위기 시기의 데이터를 기반으로 이번 코로나 시기를 예측하면 과거 데이터도 필요하고
   4. 코로나 이전의 데이터를 기반으로 한다면 2017~2018년 이전 자료를 사용할 것으로 생각됨.
   5. 한계기업으로 분류되는 재무 데이터의 영향이 당해가 아니라 시차가 존재할 수 있기 때문에 정의한 한계기업 기간에 추가적으로 기간을 덧붙여서 조사 가능.

기간: 금융위기 2008~2009 2개년, 닷컴버블 기간: 2001~2002, imf: 1997~1998

1. 대상 기업
   1. 상장 및 외감기업
   2. 규모에 따른 데이터 세트 분리
   3. 상장, 비상장 외감 기업 분리
   4. 업종 분리

외감기업이 6만개 가량 나오는데 코스피, 코스닥, 코넥스 포함하면 더 많을 것으로 생각된다.

1. 목표
   1. 1차: 기업 데이터를 기반으로 한계기업 상태를 벗어나는지 확인한다.
   2. 2차: 한계기업 상태를 벗어난 기업들을 대상으로 얘네들이 정상상태를 유지할 수 있는지 확인한다.
2. 분석 방법
   1. 피처 선정: 각종 재무 데이터
   2. 타겟: 각 데이터별 3년차 이자보상배율(‘가정’) 기반 한계기업 여부
   3. 데이터 전처리
      1. 결측치

제거

* + 1. 이상치

윈저라이징, quantiledistributon, 양쪽 0.5% 삭제

* + 1. 스케일링

Quantiledistributon

Minmax

Log

Standardscale

* + 1. 데이터 검정

정규성 검정

다중 공선성

T 검정

변수 선택

* 1. 모델
     1. 로지스틱 회귀
     2. 회귀 모형
     3. 분류 모형
        1. SVM
        2. KNN
        3. Random forest
        4. ANN, DNN
        5. RNN: LSTM, GRU
        6. CNN
        7. XGBoostClassifier, catboost, lightgbm
  2. 모델 평가
     1. 혼동행렬
     2. 재현율, 정밀도, 정확도, f1-score
     3. 모델별 선정 변수와 재무비율을 기반으로 코로나 기간에 가져와서 예측을 진행한다. -> 회생 가능성이 있는 애들을 분리한다.

예측 결과가 나와요.

1차: 과거 데이터를 기반으로 한계기업 중 회생 가능성이 있는 기업들을 분리하는 예측 모형을 만들어 내고 이것을 현재 시점에 적용해 결과물을 보여준다.

이때, 회생 가능성이 있는 기업들이 어떤 재무비율의 특징을 갖는지 설명하면 좋을 것 같다.

2차: 회생 가능성이 있는 기업들 중 정상화 상태를 유지하는 기업과 유지하지 못하는 기업들을 분류한다.

정상 – 한계기업 ------ 정상회복 -> 여기 피처

-정상기업 ㄴ 한계기업 유지

비정상 – 한계기업 ------ 정상회복 -> 여기 피처랑 차이 있을 것으로 생각되니

-정상기업 ㄴ 한계기업 유지

정상 --- 상폐종목 -🡪

비정상 --- 상폐종목

금융기관에 경제 충격 시기에 기업 여신 평가에 있어서 다른 기준을 적용할 필요성을 제언