Title, Authors

“기업 재무 정보를 이용한 ESG 등급 예측 모델”

Overall ESG grades-classifier model based on Corporation’s financial statements

Koo Minkyu, Park Minseung

**Introduction**

* What your topic is: 전자공시 API에서 제공하는 기업의 재무 데이터를 통해서 ESG 성과 예측 모델을 구축
  + To classify domestic corporation’s ESG performance grades based on financial statements.
  + ESG 경영은 투명한 지배 구조, 기업의 사회적 책임, 그리고 기후위기로 인한 사회적 요구를 반영하는 지속 가능한 사업
  + 동시에 많은 투자사와 국가 기관은 기업의 ESG 활동에 따라 앞으로의 투자를 결정하는 ESG 책임투자를 진행 중
  + Korean Institute of Corporate governance and sustainability는 환경, 사회, 일반상장사 지배구조, 금융사 지배구조 영역별 등급과 ESG 통합 등급이 부여
  + 기업은 자사의 재무적 상황에 맞게 ESG 전략을 수립하고 경영활동을 수립하기에 재무제표를 살펴보아 다음년도 ESG 통합 등급을 예측하고 ESG 책임 투자를 실행하는 투자사에 인사이트를 제공할 수 있음
  + 또한, 기준원의 ESG 등급이 기업의 ESG 활동 여력을 나타낸다고 봤을 때 ESG 활동을 지지하는 재무 지표는 무엇인지를 확인할 수 있음
* A brief literature review to summarize the state-of-art research on your topic
  + On February 2023, the authors built an algorithm that classifies corporate’s ESG grades. They collected not only the numeric values representing corporate’s financial statement but also ESG overall graders such as S, A+, A, B+, B, C, and D offered by Korean Institute of Corporate governance and sustainability.
  + By considering past data from 2019 to 2021, they could obtain 1,780 rows with 11 columns where each row represents one of domestic corporation.
  + 한국 ESG 기준원에서 제공하는 2019년부터 2021년 기업별 ESG 통합등급 데이터 1,780건과 1,780개의 기업의 재무 데이터를 수집하여 통합등급 B+ 이상과 B 이하로 구분하는 이진 분류 모델을 만듦
  + 재무구조, 성장성, 지분 구조 및 기타 영역을 파악하는 11개 변수를 선정하고 ESG 기준원의 등급을 타겟 변수로 삼아 심층신경망 모델을 구축함
* Overall plan
  + 참조 논문처럼 사용한 칼럼과 데이터를 사용
  + 연도별 데이터를 추가로 수집하여 데이터 양을 늘릴 계획
  + 다른 논문을 확인하여 추가로 고려할 칼럼을 추출하여 ESG 통합 등급에 영향을 끼치는 변수의 중요성을 알아봄

**Teammate and work division**

* Plan along with activities of each group member
  + Project Proposal: 민규
  + Planning & Topic selection: 모두
  + Dataset preparation (API request and Making dataframe): 민승
  + Data Preprocessing: 민규
  + Model selection & training: All
  + Evaluation & Parameter tuning: All
* Be sure to specify how each team member will contribute to the project

**Problem Statement (정리)**

* Be sure to clearly state the research questions
  + problem definition and dataset
    - Broader approach about ML algorithm that predicts future ESG overall grades based on corporation’s financial statement
    - Dataset: financial statements and ESG grades table
    - One is collected by the library and the tables was being crawled
* 가정 1. 재무구조 및 성장성이 높으면 높은 ESG 등급을 받는다
* 가정 2. 임직원 성장률과 사외 이사 비율이 크게 작용할 듯싶다
* Expected results and evaluation

Technical Approach

* **multiclass classification task) Logistic Regression, Decision Tree, Random Forest, NN**

모델 성능뿐만 아니라 특성의 중요도를 쉽게 확인할 수 있다는 알고리즘 필요

* results and evaluation: 분류 모델 -> 다중 분류의 정확도, 전반적인 성능을 확인하고자 F1 score와 AUC-ROC 곡선 확인
* Discuss any limitations of your proposed approach
  + Company 수 자체가 몇 천 개밖에 안 되고, 연도별로 독립적으로 추출해도 배타적이라고 볼 수 없음