Project title

Data-driven ESG scoring with NLP

Project description

1. Problem statement
   Financial institutions (FIs) are facing increasing pressure on Environmental, Social, and Governance (ESG) disclosure, where they are commonly required to release a yearly report. However, there is no regulatory guideline or standardised framework for ESG reporting, causing information to be communicated in various manners among different reports. As such, it is exhausting for third-party agencies (e.g., MSCI) to manually consolidate information and create benchmarked ESG metrics, and time-consuming for individual investors to go through these reports and make sound investment decisions.

2. Objectives
   The aim of our final year project is to streamline the data extraction process from ESG disclosure reports of various FIs and use the collected information to analyse, display and compare the financial performance of these FIs.

3. Deliverables
   Firstly, we intend to utilise natural language processing (NLP) techniques to convert the ESG reports of chosen FIs from their raw PDF formats into well-defined sentences, and use Latent Dirichlet Allocation (LDA) model to classify ESG statements and summarise key ESG initiatives of these FIs. We would then generate an ESG score for each FI based on the data gathered from the NLP process. We can check the accuracy of our ESG scoring model by comparing it against the ESG Risk Ratings on Yahoo Finance and using it to further improve our model.

   The final deliverable would be in the form of a webpage application. End users can obtain ESG scores of their desired institution as well as the institution’s financial data pulled from Yahoo Finance API, showing a holistic overview of the company’s performance. Utilising visualisation techniques, we would also compare how different FIs focus on different ESG initiatives, and display a ranking of ESG performance against other FIs in the same industry.

   With this application, we aim to offer ESG-as-a-service to financial investors so they can better invest in sustainable and responsible finance.

Grouping information

1. Ding, Ai Wei (3035719741) – BASc(FinTech)
2. Leong, Winky (3035664374) – BASc(FinTech)
3. Tanoto, Nicholas (3035667041) – BASc(FinTech)

Proposed supervisor

Professor Yiu, Siu Ming (smyiu@cs.hku.hk) – obtained consent for supervision