PICTET
1805
Asset Management

A case study for ESG data scientist:

Which country will you invest?

DESCRIPTION

Imaging you have CHF 10 million and you want to invest in sovereign bonds, which country would you invest in?

As an ESG-enthusiast, you first would like to evaluate countries based on ESG criteria. You can find a list of country-level data from The World Bank API (https://datahelpdesk.worldbank.org/knowledgebase/topics/125589). There is also a python wrapper (https://wbdata.readthedocs.io/en/stable/) that helps you navigate and retrieve data from the api. From these data, can you choose the appropriate metrics and construct a country-level ESG rating?

In addition, you want to have a good balance between risk and return for your investment. Country risk premium can be estimated by comparing the spread on sovereign debt yields between the country and a mature market like the U.S. We have provided you the country spreads of the past 10 years for some emerging countries. As a data scientist, can you build a model to predict the spreads with the data you get from the World Bank API? How do you evaluate the performance of your model(s)?

Please feel free to add other data sources that you think are relevant to this study.

DELIVERABLE

You will present your work during an 1h session, with Q&A. The audience is composed of a mix of technical and non-technical staff. As a candidate for **data scientist**, you are expected to demonstrate your proficiency in understanding business questions, your engineering skills in retrieving and structuring relevant data from various sources, your data analysis skills, and you capabilities in training/evaluating machine learning models.

You may organise your presentation according to the following structure. Please **focus on the bold parts** as the main content of your presentation.

— Business understanding:
    - problem definition

— Data acquisition and understanding:
    - data sources and formats
    - cleaning and wrangling
    - data exploration

— **Analysis:**
    - model construction
    - model evaluation
    - interpreting results

— Deployment:
    - how to deploy the pipeline on the cloud, such as AWS.

This is just a suggestion, please feel free to change the structure in any way you feel best.

Please **send us your code (*i.e.* git repo)** one day before the interview.