■入社試験課題

(1) arxiv.org の論文をスクレイピングするコードを作成せよ (Abstractは最低限、本文も読めると最高)

(2) 任意のキーワードで arxiv.org の論文を検索できるよう拡張せよ

(3) bioarxiv.org にも対応せよ

課題（１）の疑問点

1.「abstract」の定義が「要旨（Abstract）」の意味であるか確認が必要。→ABC (Abstract Base Class - 抽象基底クラス)　を使用してのコーディングと当初、勘違いしていたので。

2.初回コード（v1.0)では、特定論文ページでクラス指定で「要旨（Abstract）」スクレイピングを実装→サイト全体でおこなう場合のループ処理？が手順が不明。ハイパーリンクで、複数テキストを生成し、サマリシートをすべきか、長文で全データを吐き出すべきかを再確認。→JupyterLabの画面が実行時に全文吐き出しだと操作性が問題ありそうなため。

3.あえてタグを外していないのと整形の仕様が不慣れため想像ができない。成果物イメージの情報を確認する。

4.Vscode 環境では、実行時エラー（初歩的な問題であるが対応の調査に時間が

　かかる可能性があり保留している。

PS C:\Users\user> conda activate base

conda : 用語 'conda' は、コマンドレット、関数、スクリプト ファイル、または操作可能なプログラムの名前として認識されません。名前

が正しく記述されていることを確認し、パスが含まれている場合はそのパスが正しいことを確認してから、再試行してください。

+ conda activate base

+ ~~~~~

+ CategoryInfo : ObjectNotFound: (conda:String) [], CommandNotFoundException

+ FullyQualifiedErrorId : CommandNotFoundException

5.上記の1ｰ4がクリアできた時点でクラス指定で本文のスクレイピングを実装。

課題（２）の疑問点

ポップアップ等で検索キーワード入力で良いかのオペレーションを確認してから着手

■ Entrance exam assignment

(1) Write code to scrape arxiv.org treatises (Abstract is at least as good as reading the text)

(2) Expand so that you can search arxiv.org articles with any keyword.

(3) Also support bioarxiv.org

Question of issue (1)

1. It is necessary to confirm whether the definition of "abstract" means "abstract". → I was initially mistaken for coding using ABC (Abstract Base Class).

2. In the initial code (v1.0), implement "Abstract" scraping by specifying the class on the specific paper page → Loop processing when performing on the entire site? But the procedure is unknown. Reconfirm whether you should generate multiple texts with hyperlinks and make a summary sheet or spit out all the data in a long sentence. → If the Jupyter Lab screen spits out the full text at runtime, there seems to be a problem with operability.

3. I can't imagine because I didn't remove the tag and I'm not used to the plastic surgery specifications. Check the information of the deliverable image.

4. In Vscode environment, run-time error (a rudimentary problem, but it takes time to investigate the response

It may take some time and is on hold.

PS C: \ Users \ user> conda activate base

conda: The term'conda' is not recognized as the name of a cmdlet, function, script file, or operable program. given names

Make sure that is written correctly, and if it contains a path, make sure that the path is correct, then try again.

+ conda activate base

+ ~~~~~

+ CategoryInfo: ObjectNotFound: (conda: String) [], CommandNotFoundException

+ FullyQualifiedErrorId: CommandNotFoundException

5. When the above 1-4 can be cleared, the text scraping is implemented by specifying the class.

Question of issue (2)

Start after confirming the operation whether it is okay to enter the search keyword in the pop-up etc.