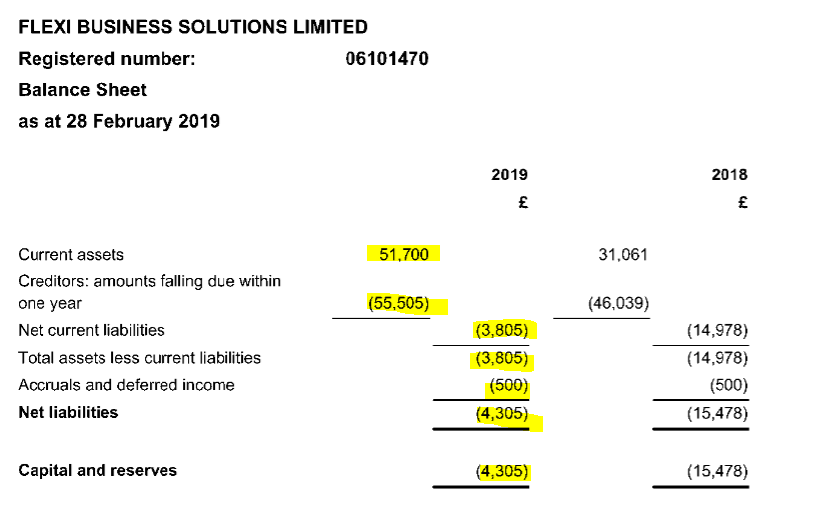
**Problem Statement Focus:** To check for NLP and Python capability to extract data elements from the financial statement documents.

**Data to be provided –**

* **Training & Development Set -** 500 Text files (OCR output of images)

**Sample Data to be Extracted**

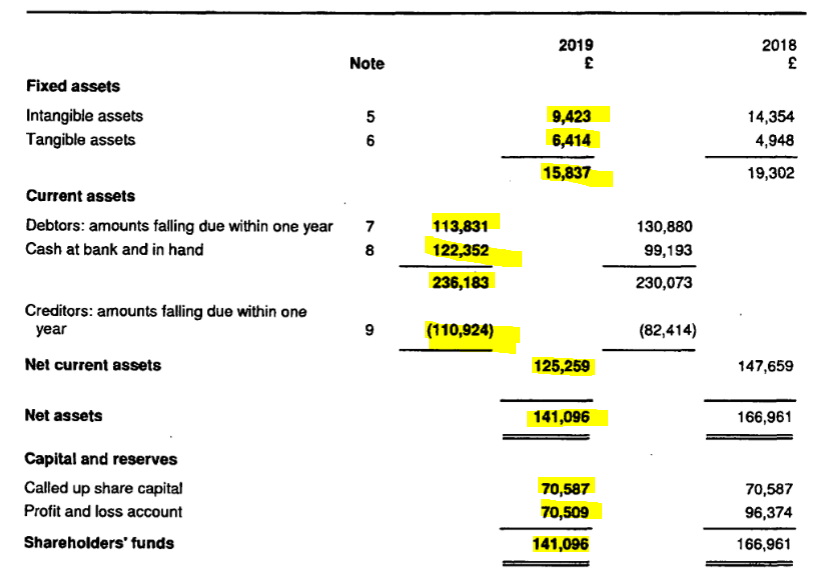
Highlighted text on the image is required to be extracted



Fields to be extracted are:

* Current Assets : 51,700 -> 51700
* Creditors: amounts falling due within one year : 55,505 -> 55505
* Net current liabilities : (3,805) -> 3805
* Total assets less current liabilities : (3,805) -> 3805
* Accruals and deferred income : (500) -> -500
* Net liabilities : (4,305) -> -4305
* Capital and reserves : (4,305) -> -4305

**Results file entry:** {“Current assets":"51700","Creditors:amounts falling due within one year":"-55505","Net current liabilities":"-3805","Total assets less current liabilities":"-3805","Accurals and deferred income":"-500","Net liabilities":"-4305","Capital and reserves":"-4305"}



Fields to be extracted are:

* Fixed asset :  15837
* Intangible assets : 9423
* Tangible assets : 6414
* “nan” : 15,837 -> 15837
* Current Assets : 236,183 -> 236183
* Debtors: amounts falling due within one year : 113,831 -> 113831
* Cash at bank and in hand : 122,352 -> 122352
* Creditors: amounts falling due within one year : (110,924) -> -110924
* Net current assets : 125,259 → 125259
* Net assets : 141,096 → 141096
* Capital and reserves: -
* Called up share capital : 70,587 -> 70587
* Profit and loss accounts 70,509 -> 70509
* Shareholders funds : 141,096 -> 141096

Results file entry: {"nan": "236183", "Current assets": "nan", "Creditors: amounts falling due within one year": "{110924)", "Shareholders funds": "141096", "Intangible assets": "9423", "Fixed assets": "nan", "Net assets": "141096", "Tangible assets": "6414", "Called up share capital": "70587", "Capital and reserves": "nan", "Net current assets": "125259", "Profit and loss account": "70509", "Cash at bank and in hand": "122352", "Debtors: amounts falling due within one year": "113831"}

Please note, data only for year 2019 is to be extracted.