In [1]:
```python
from bs4 import BeautifulSoup
import re
import unicodedata
import xml.etree.cElementTree as ET
import pandas as pd
```

Code for extracting all paragraphs from xml paper

In [2]:
```python
repo = []
count = 0
for i in range(0, 4128):
    try:
        tree = ET.ElementTree(file = "G:\\mysearch_elsevier_xml_papers\\mysearch_els_xml_" + str(i) + ".xml")
        paper_detail = []

        root = tree.getroot()
        for roots in root[0]:
            if roots.tag == '{http://prismstandard.org/namespaces/basic/2.0/}doi': # for extracting doi
                text = roots.text
                paper_detail.append(text)

            if roots.tag == '{http://purl.org/dc/elements/1.1/}title': # for extracting title
                text = ET.tostring(roots, encoding='unicode', method='text' )
                text = re.sub(r"[\n]( )*", "", text)
                text = unicodedata.normalize("NFKD",text)
                text = text.strip()
                paper_detail.append(text)


            if roots.tag == '{http://purl.org/dc/elements/1.1/}description': # for extracting abstract
                text = ET.tostring(roots, encoding='unicode', method='text')
                text = re.sub(r"[\n]( )*", "", text)
                text = unicodedata.normalize("NFKD",text)
                text = text.strip()
                paper_detail.append(text)
#                 print(text, end = "*********************\n**************")

        para = []
        for p in tree.iter():                  # paragraphs extraction starts here
            if(p.tag =='{http://www.elsevier.com/xml/common/dtd}para'):
                try:
                    text = ET.tostring(p, encoding='unicode', method='text' )
                    text = re.sub(r"[\n]( )*", "", text)
                    text = unicodedata.normalize("NFKD",text)
                    text = text.strip()
                    para.append(text)
                except:
                    continue
        paper_detail.append(para)
        paper_detail.append(len(para)) # appending total number of paragraphs
        count = count + len(para)  # counting total number of paragraphs

        repo.append(paper_detail)
    except:
        continue
print(len(repo), "Number of para is ", count)
# repo
```

```
4124 Number of para is  88989
```

In [3]:
```python
df = pd.DataFrame(repo)
df.index.names = ["Serial_No"]
df.columns = ["DOI","Title","Astract","Para_list" , "Para_count"]
pd.set_option('display.max.rows',8580)
df.head(50)
```

Out[3]:

Serial_No	DOI	Title	Astract	Para_list	Para_count
0	10.1016/j.spmi.2021.106983	Structural behaviour of BiFeO3/SrRuO3 superlat...	Epitaxial BiFeO3/SrRuO3 superlattices have bee...	[BiFeO3 (BFO) is under intense investigation b...	19
1	10.1016/j.spmi.2018.01.027	Chemically stabilized epitaxial wurtzite-BN th...	We report on the chemically stabilized epitaxi...	[Boron Nitride (BN) exhibits numerous allotrop...	19
2	10.1016/j.spmi.2016.11.031	High quality interconnected core/shell ZnO nan...	We report the production of vertically aligned...	[Interconnected architectures of ZnO nanostruc...	17
3	10.1016/j.jcis.2015.01.088	Strontium and zoledronate hydroxyapatites grad...	Both strontium and zoledronate (ZOL) are known...	[Orthopaedic implants require materials with r...	38
4	10.1016/j.actamat.2021.117257	Rheology of amorphous olivine thin films chara...	The rheological properties of amorphous olivin...	[Olivine, a silicate with composition (Mg,Fe)2...	48
5	10.1016/j.actamat.2021.116955	Novel class of nanostructured metallic glass f...	A novel class of nanostructured Zr50Cu50 (%at....	[Bulk metallic glasses (BMGs) are characterize...	32
6	10.1016/j.actamat.2020.09.035	Differences in Sb2Te3 growth by pulsed laser a...	High quality van der Waals chalcogenides are i...	[Sb2Te3, Bi2Te3, and Bi2Se3 are typical van de...	41
7	10.1016/j.tsf.2015.09.060	Effect of self-grown seed layer on thermoelect...		[ZnO is an n-type semiconductor with versatile...	15
8	10.1016/j.jlumin.2015.08.031	Er3+-doped fluorotellurite thin film glasses w...	Transparent oxyfluoride tellurite thin film gl...	[The age of all-optical metropolitan and local...	33
9	10.1016/j.ssi.2015.10.010	Li+ conduction in Li–Nb–O films deposited by a...	We fabricated amorphous Li–Nb–O films by a sol...	[Many researchers have devoted significant eff...	27
10	10.1016/j.tsf.2015.12.066	Preparation of TaO2 thin films using NbO2 temp...	TaO2 thin films were prepared using rutile-typ...	[TaO2 has a rutile-type structure with a tetra...	25
11	10.1016/j.jpowsour.2015.09.097	Reversible lithium intercalation in a lithium-...	Li2RuO3 (001) films with a lithium-rich layere...	[All-solid-state lithium batteries are attract...	14
12	10.1016/j.nima.2015.09.064	Nanomechanical and electrical properties of Nb...	We report a design of photocathode, which comb...	[The use of photoinjectors based on supercondu...	17
13	10.1016/j.nima.2015.09.020	Development and testing of cost-effective, 6cm...	Micro-channel plate (MCP)-based photodetectors...	[Photomultiplier tubes (PMTs) [1] are current-...	50
14	10.1016/j.jnucmat.2014.10.016	Laser cleaning of diagnostic mirrors from toka...	This paper presents a laboratory-scale experim...	[Metallic First Mirrors (FMs) will be crucial ...	14
15	10.1016/j.ssi.2015.02.005	Crystallization process of perovskite type oxi...	Crystallization process in non-heating pulsed ...	[Perovskite oxide based proton conducting cera...	26
16	10.1016/j.solmat.2015.06.004	Molecular beam and pulsed laser deposition of ...	We have investigated the structural and optica...	[An intermediate band material has at least on...	47
17	10.1016/j.jcrysgro.2015.01.022	Structural, optical and electrical properties ...	Sb2Te3 films were deposited using pulsed laser...	[Sb2Te3 is a V–VI compound semiconductor, and ...	11
18	10.1016/j.jmmm.2014.11.055	Interrelation of epitaxial strain and oxygen d...	The interrelation between the epitaxial strain...	[Infinite layer manganites of the type AMnO3, ...	18
19	10.1016/j.tsf.2015.09.066	Tuning electrical properties of hierarchically...	Large surface area, 3D structured transparent ...	[Aluminum doped Zinc Oxide (AZO) is an inexpen...	27
20	10.1016/j.astropartphys.2014.12.003	Thermo-acoustic sound generation in the intera...	The generation of hydrodynamic radiation in in...	[In 1957 G.A. Askaryan pointed out that ionisa...	55
21	10.1016/j.jnucmat.2014.11.046	Epitaxial Fe/Y2O3 interfaces as a model system...	The fundamental mechanisms underlying the supe...	[Next-generation fission and future fusion rea...	22
22	10.1016/j.physc.2014.07.007	Structural and electronic properties of epitax...	Epitaxial bilayers of the high-temperature-sup...	[Superlattices and heterostructures of transit...	13
23	10.1016/j.ssi.2013.09.040	Hetero-epitaxial growth of Li0.17La0.61TiO3 so...	A Li0.17La0.61TiO3(111)/LiMn2O4(111) solid ele...	[Perovskite-type lithium lanthanum titanates, ...	12
24	10.1016/j.ssi.2013.09.054	Fabrication of thin-film lithium batteries wit...	All-solid-state thin-film batteries with a 5-V...	[All-solid-state thin-film batteries have attr...	22
25	10.1016/j.nimb.2014.02.100	Swift heavy ion induced nano-dimensional phase...	Pulsed laser deposited 60nm thin film of homog...	[Materials composed of dispersed soft metal na...	12
26	10.1016/j.nimb.2014.02.104	Channelling study of La1−xSrxCoO3 films on dif...	The cobalt oxide system LaCoO3 and its Sr-dope...	[LaCoO3 perovskite and its Sr-doped derivative...	19
27	10.1016/j.nima.2014.03.032	Non-conventional photocathodes based on Cu thi...	Copper (Cu) thin films were deposited on yttri...	[Metal-based photocathodes are being used in t...	24
28	10.1016/j.cap.2014.10.016	Electrical and electronic properties of nitrog...	Nitrogen-doped amorphous carbon thin films (a-...	[Amorphous carbon (a-C) thin film has attracte...	8
29	10.1016/j.jmmm.2014.06.038	Thickness dependence of dynamic and static mag...	We present a comprehensive study of the thickn...	[Manganites have attracted large attention due...	33
30	10.1016/j.jmmm.2009.10.032	Growth mode, magnetic and magneto-optical prop...	The growth mode, magnetic and magneto-optical ...	[Owing to its ability to essentially preserve ...	16
31	10.1016/j.colsurfa.2009.09.039	Synthesis and rheological properties of cation...	In this paper we report our new approach to sy...	[In geological drilling, drilling fluids are i...	26
32	10.1016/j.apsusc.2009.06.106	The effect of relative plasma plume delay on t...	We report the effects of relative time delay o...	[Pulsed laser deposition (PLD) has proven to b...	25
33	10.1016/j.apsusc.2008.08.050	Time resolved Nomarski interferometery of lase...	We report results from optical interferometric...	[PLD is a well-established method in materials...	13
34	10.1016/j.apsusc.2007.11.016	A comparative study of gadolinium gallium garn...	The growth of epitaxial Nd:Gd3Ga5O12 (GGG) on ...	[The study of nanosecond laser ablation and pu...	14
35	10.1016/j.apsusc.2007.04.089	Parallel syntheses and thermoelectric properti...	Thermoelectric properties of single crystallin...	[Combinatorial chemistry has been developed in...	10
36	10.1016/j.apsusc.2007.07.122	Magnetic and structural study of Cu-doped TiO2...	Transparent pure and Cu-doped (2.5, 5 and 10at...	[A variety of doped-semiconducting materials, ...	11
37	10.1016/j.physc.2007.03.104	Microwave properties of YBa2Cu3O7−δfilms with ...	We present measurements of the microwave compl...	[Various techniques have been exploited in ord...	3
38	10.1016/j.tsf.2006.11.075	Ferroelectric Bi3.25La0.75Ti3O12 thin films on...	Strontium ruthenate and Bi3.25La0.75Ti3O12 (BL...	[Today, the main potential application field f...	15
39	10.1016/j.surfcoat.2006.08.017	UHV arc for high quality film deposition	The vacuum arc is a well-known technique for p...	[Vacuum arc is one of the oldest techniques us...	22
40	10.1016/j.apsusc.2005.02.125	Magnetic and spectroscopic characteristics of ...	We report on the observation of room-temperatu...	[The observation of ferromagnetic transition i...	11
41	10.1016/j.apsusc.2003.09.039	Thick film growth of high optical quality low ...	Thick film growth of high optical quality Nd:G...	[Pulsed laser deposition (PLD) is well establi...	24
42	10.1016/j.apsusc.2003.07.005	High throughput characterization of the optica...	Compositionally graded combinatorial films hav...	[The use of combinatorial chemistry techniques...	11
43	10.1016/S1369-7021(04)00287-1	ZnO: growth, doping & processing	A review is given here of recent results in de...	[Recent improvements in the control of backgro...	15
44	10.1016/S0040-6090(01)01178-6	Characterization of polycrystalline Cu(In,Ga)T...	Thin films of the chalcopyrite compound CuGaXI...	[CuInTe2 and CuGaTe2 ternary compounds are dir...	17
45	10.1016/S0921-4534(01)00398-7	Superconducting magnesium diboride films with ...	Thin superconducting films of magnesium dibori...	[The recent discovery of superconductivity at ...	17
46	10.1016/S0927-0248(97)00223-7	Electrochromic lithium nickel oxide by pulsed ...	Thin films of lithium nickel oxide were deposi...	[The layered form of lithium nickel oxide LiNi...	21
47	10.1016/S0927-0248(97)00222-5	Optical indices of lithiated electrochromic ox...	Optical indices have been determined for thin ...	[The complex refractive index of electrochromi...	15
48	10.1016/S0039-6028(97)01072-8	Carbon-based nanostructured materials via clus...	The use of clusters as elemental building bloc...	[In recent years, the nanoscale materials synt...	18
49	10.1016/S0169-4332(97)00774-5	Low-temperature growth of YBCO thin films by p...	The effect of ambient gas on the preparation o...	[It is very important for growth of high-quali...	13

In [10]:
```python
len(df['Para_list'])
```

Out[10]:
```
4124
```

In [11]:
```python
paras = []
for para_list in df['Para_list']:
    for paragraphs in para_list:
        paras.append(paragraphs)


len(paras)
```

Out[11]:
```
88989
```

In [12]:
```python
para_df = pd.DataFrame(paras)
para_df.index.names = ['Serial_No']
para_df.columns = ["Paragraph"]
para_df.head(10)
```

Out[12]:

Serial_No	Paragraph
0	BiFeO3 (BFO) is under intense investigation be...
1	The SLs were grown by pulsed laser deposition ...
2	Table 1below presents the characteristics of t...
3	Fig. 1presents the X-ray diffraction pattern i...
4	In order to get access to the out of plane lat...
5	The overall correspondence is good considering...
6	We observe on Table 2 an increase of the out o...
7	To confirm the thicknesses deduced from the ab...
8	Satellite maxima are well pronounced at low an...
9	To better investigate the structural evolution...

In [6]:
```python
df.to_csv("G:\\paper_to_para\\xml_4128_para.csv",index = False )
```

In [13]:
```python
para_df.to_csv("G:\\paper_to_para\\xml_4128_only_para.csv",index = False )
```

code for publisherless xml Paper

In [23]:
```python
repo_1 = []
for i in range(0, 2292):
    try:
        tree = ET.ElementTree(file = "G:\\publisherless_elsvier_xml\\xml_" + str(i) + ".xml")
        paper_detail = []

        root = tree.getroot()
        for roots in root[0]:
            if roots.tag == '{http://prismstandard.org/namespaces/basic/2.0/}doi': # for extracting doi
                text = roots.text
                paper_detail.append(text)

            if roots.tag == '{http://purl.org/dc/elements/1.1/}title': # for extracting title
                text = ET.tostring(roots, encoding='unicode', method='text' )
                text = re.sub(r"[\n]( )*", "", text)
                text = unicodedata.normalize("NFKD",text)
                text = text.strip()
                paper_detail.append(text)


            if roots.tag == '{http://purl.org/dc/elements/1.1/}description': # for extracting abstract
                text = ET.tostring(roots, encoding='unicode', method='text')
                text = re.sub(r"[\n]( )*", "", text)
                text = unicodedata.normalize("NFKD",text)
                text = text.strip()
                paper_detail.append(text)
#                 print(text, end = "*********************\n**************")

        para = []
        for p in tree.iter():                  # paragraphs extraction starts here
            if(p.tag =='{http://www.elsevier.com/xml/common/dtd}para'):
                try:
                    text = ET.tostring(p, encoding='unicode', method='text' )
                    text = re.sub(r"[\n]( )*", "", text)
                    text = unicodedata.normalize("NFKD",text)
                    text = text.strip()
                    para.append(text)
                except:
                    continue
        paper_detail.append(para)
        paper_detail.append(len(para)) # appending total number of paragraphs
        repo_1.append(paper_detail)

        count = count + len(para)  # counting total number of paragraphs

        repo.append(paper_detail)
    except:
        continue
print(len(repo), "Number of para is ", count)
```

```
6415 Number of para is  134236
```

In [31]:
```python
df2 = pd.DataFrame(repo_1)
df2.index.names = ["Serial_No"]
df2.columns = ["DOI","Title","Astract","Para_list" , "Para_count"]
pd.set_option('display.max.rows',8580)
df2.head(5)
```

Out[31]:

Serial_No	DOI	Title	Astract	Para_list	Para_count
0	10.1016/j.carbon.2014.01.061	Giant photoconductivity induced by plasmonic C...	A giant photoconductivity was observed in hete...	[The family of a-C films is a potential semico...	21
1	10.1016/j.materresbull.2014.03.005	The shift of optical band gap in W-doped ZnO w...	Tungsten-doped (W-doped) zinc oxide (ZnO) nano...	[Zinc oxide (ZnO) nanomaterials are promising ...	15
2	10.1016/j.nima.2014.02.004	State-of-the-art Pb photocathodes deposited by...	In this article we present and discuss the cur...	[The development of metallic photocathodes has...	29
3	10.1016/j.jallcom.2014.01.160	Electrical transport behavior of lead-free 0.5...	The temperature and voltage-polarity dependent...	[The advances in the thin film growth technolo...	14
4	10.1016/j.bios.2013.11.015	Inducing electrocatalytic functionality in ZnO...	A third generation uric acid biosensor has bee...	[Uric acid is the end metabolic product of pur...	38

In [26]:
```python
df2.to_csv("G:\\paper_to_para\\xml_2292_para.csv",index = False )
```

In [27]:
```python
paras = []
for para_list in df2['Para_list']:
    for paragraphs in para_list:
        paras.append(paragraphs)


len(paras)
```

Out[27]:
```
45247
```

In [28]:
```python
para_df = pd.DataFrame(paras)
para_df.index.names = ['Serial_No']
para_df.columns = ["Paragraph"]
para_df.head(10)
```

Out[28]:

Serial_No	Paragraph
0	The family of a-C films is a potential semicon...
1	Here, we mainly investigate the photoconductiv...
2	We demonstrate experimentally that Co nanopart...
3	Co–C and pure a-C films were deposited by the ...
4	Scanning electron microscope (SEM) was used to...
5	Silver surface electrodes were evaporated on C...
6	The magnitude of photoconductivity is defined ...
7	PLD was used to grow p-type Co–C films on n-ty...
8	Raman spectroscopy was adopted to characterize...
9	I–V characteristics of the Co–C/Si heterostruc...

In [29]:
```python
para_df.to_csv("G:\\paper_to_para\\xml_2292_only_para.csv",index = False )
```

Extracting all para from both xml papers

In [32]:
```python
df3 = pd.DataFrame(repo)
df3.index.names = ["Serial_No"]
df3.columns = ["DOI","Title","Astract","Para_list" , "Para_count"]
pd.set_option('display.max.rows',8580)
len(df3['DOI'])
```

Out[32]:
```
6415
```

In [33]:
```python
df3.to_csv("G:\\paper_to_para\\all_xml_para.csv",index = False )
```

In [34]:
```python
paras = []
for para_list in df3['Para_list']:
    for paragraphs in para_list:
        paras.append(paragraphs)


len(paras)
```

Out[34]:
```
134236
```

In [35]:
```python
para_df = pd.DataFrame(paras)
para_df.index.names = ['Serial_No']
para_df.columns = ["Paragraph"]
para_df.head(10)
```

Out[35]:

Serial_No	Paragraph
0	BiFeO3 (BFO) is under intense investigation be...
1	The SLs were grown by pulsed laser deposition ...
2	Table 1below presents the characteristics of t...
3	Fig. 1presents the X-ray diffraction pattern i...
4	In order to get access to the out of plane lat...
5	The overall correspondence is good considering...
6	We observe on Table 2 an increase of the out o...
7	To confirm the thicknesses deduced from the ab...
8	Satellite maxima are well pronounced at low an...
9	To better investigate the structural evolution...

In [36]:
```python
para_df.to_csv("G:\\paper_to_para\\all_xml_only_para.csv",index = False )
```

Reading Extracted paragraphs

In [45]:
```python
xml = pd.read_csv("G:\\paper_to_para\\all_xml_para.csv" ,converters = {"Para_list": eval}  )
# xml.set_index("Serial_No" , inplace = True)
xml.head(5)
```

Out[45]:

	DOI	Title	Astract	Para_list	Para_count
0	10.1016/j.spmi.2021.106983	Structural behaviour of BiFeO3/SrRuO3 superlat...	Epitaxial BiFeO3/SrRuO3 superlattices have bee...	[BiFeO3 (BFO) is under intense investigation b...	19
1	10.1016/j.spmi.2018.01.027	Chemically stabilized epitaxial wurtzite-BN th...	We report on the chemically stabilized epitaxi...	[Boron Nitride (BN) exhibits numerous allotrop...	19
2	10.1016/j.spmi.2016.11.031	High quality interconnected core/shell ZnO nan...	We report the production of vertically aligned...	[Interconnected architectures of ZnO nanostruc...	17
3	10.1016/j.jcis.2015.01.088	Strontium and zoledronate hydroxyapatites grad...	Both strontium and zoledronate (ZOL) are known...	[Orthopaedic implants require materials with r...	38
4	10.1016/j.actamat.2021.117257	Rheology of amorphous olivine thin films chara...	The rheological properties of amorphous olivin...	[Olivine, a silicate with composition (Mg,Fe)2...	48

In [52]:
```python
html_1 = pd.read_csv("G:\\paper_to_para\\html_1007_para.csv" ,converters = {"Para_list": eval}  )
html_1.shape
```

Out[52]:
```
(987, 5)
```

In [51]:
```python
html_2 = pd.read_csv("G:\\paper_to_para\\random_para.csv" ,converters = {"Para_list": eval}  )
html_2.shape
```

Out[51]:
```
(1187, 5)
```

In [53]:
```python
df_combo = pd.concat([xml , html_1 , html_2] , ignore_index= True)
df_combo.head(15)
```

Out[53]:

	DOI	Title	Astract	Para_list	Para_count
0	10.1016/j.spmi.2021.106983	Structural behaviour of BiFeO3/SrRuO3 superlat...	Epitaxial BiFeO3/SrRuO3 superlattices have bee...	[BiFeO3 (BFO) is under intense investigation b...	19
1	10.1016/j.spmi.2018.01.027	Chemically stabilized epitaxial wurtzite-BN th...	We report on the chemically stabilized epitaxi...	[Boron Nitride (BN) exhibits numerous allotrop...	19
2	10.1016/j.spmi.2016.11.031	High quality interconnected core/shell ZnO nan...	We report the production of vertically aligned...	[Interconnected architectures of ZnO nanostruc...	17
3	10.1016/j.jcis.2015.01.088	Strontium and zoledronate hydroxyapatites grad...	Both strontium and zoledronate (ZOL) are known...	[Orthopaedic implants require materials with r...	38
4	10.1016/j.actamat.2021.117257	Rheology of amorphous olivine thin films chara...	The rheological properties of amorphous olivin...	[Olivine, a silicate with composition (Mg,Fe)2...	48
5	10.1016/j.actamat.2021.116955	Novel class of nanostructured metallic glass f...	A novel class of nanostructured Zr50Cu50 (%at....	[Bulk metallic glasses (BMGs) are characterize...	32
6	10.1016/j.actamat.2020.09.035	Differences in Sb2Te3 growth by pulsed laser a...	High quality van der Waals chalcogenides are i...	[Sb2Te3, Bi2Te3, and Bi2Se3 are typical van de...	41
7	10.1016/j.tsf.2015.09.060	Effect of self-grown seed layer on thermoelect...	NaN	[ZnO is an n-type semiconductor with versatile...	15
8	10.1016/j.jlumin.2015.08.031	Er3+-doped fluorotellurite thin film glasses w...	Transparent oxyfluoride tellurite thin film gl...	[The age of all-optical metropolitan and local...	33
9	10.1016/j.ssi.2015.10.010	Li+ conduction in Li–Nb–O films deposited by a...	We fabricated amorphous Li–Nb–O films by a sol...	[Many researchers have devoted significant eff...	27
10	10.1016/j.tsf.2015.12.066	Preparation of TaO2 thin films using NbO2 temp...	TaO2 thin films were prepared using rutile-typ...	[TaO2 has a rutile-type structure with a tetra...	25
11	10.1016/j.jpowsour.2015.09.097	Reversible lithium intercalation in a lithium-...	Li2RuO3 (001) films with a lithium-rich layere...	[All-solid-state lithium batteries are attract...	14
12	10.1016/j.nima.2015.09.064	Nanomechanical and electrical properties of Nb...	We report a design of photocathode, which comb...	[The use of photoinjectors based on supercondu...	17
13	10.1016/j.nima.2015.09.020	Development and testing of cost-effective, 6cm...	Micro-channel plate (MCP)-based photodetectors...	[Photomultiplier tubes (PMTs) [1] are current-...	50
14	10.1016/j.jnucmat.2014.10.016	Laser cleaning of diagnostic mirrors from toka...	This paper presents a laboratory-scale experim...	[Metallic First Mirrors (FMs) will be crucial ...	14

In [54]:
```python
df_combo.shape
```

Out[54]:
```
(8589, 5)
```

In [55]:
```python
paras = []
for para_list in df_combo['Para_list']:
    for paragraphs in para_list:
        paras.append(paragraphs)


len(paras)
```

Out[55]:
```
163228
```

In [56]:
```python
para_df = pd.DataFrame(paras)
para_df.index.names = ['Serial_No']
para_df.columns = ["Paragraph"]
para_df.head(10)
```

Out[56]:

Serial_No	Paragraph
0	BiFeO3 (BFO) is under intense investigation be...
1	The SLs were grown by pulsed laser deposition ...
2	Table 1below presents the characteristics of t...
3	Fig. 1presents the X-ray diffraction pattern i...
4	In order to get access to the out of plane lat...
5	The overall correspondence is good considering...
6	We observe on Table 2 an increase of the out o...
7	To confirm the thicknesses deduced from the ab...
8	Satellite maxima are well pronounced at low an...
9	To better investigate the structural evolution...

In [57]:
```python
para_df.to_csv("G:\\paper_to_para\\combo_only_para.csv",index = False )
```