All NeurIPS (NIPS) Papers

MVP

- Files info :

```
papers_data.info
```

```
<bound method DataFrame.info of       source_id  year                                              title  \
0            27  1987                         Bit-Serial Neural Networks
1            63  1987                       Connectivity Versus Entropy
2            60  1987       The Hopfield Model with Multi-Level Neurons
3            59  1987                                How Neural Nets Work
4            69  1987  Spatial Organization of Neural Networks: A Pro...
...         ...   ...                                                ...
9675       5452  2019  Discrete Object Generation with Reversible Ind...
9676       4799  2019  Adaptively Aligned Image Captioning via Adapti...
9677       1827  2019        Fully Dynamic Consistent Facility Location
9678       8693  2019     Efficient Rematerialization for Deep Networks
9679       2302  2019  Flow-based Image-to-Image Translation with Fea...

                                             abstract  \
0                                                 NaN
1                                                 NaN
2                                                 NaN
3                                                 NaN
4                                                 NaN
...                                               ...
9675  The success of generative modeling in continuo...
9676  Recent neural models for image captioning usua...
9677  We consider classic clustering problems in ful...
9678  When training complex neural networks, memory ...
9679  Learning non-deterministic dynamics and intrin...

                                              full_text
0     573 \n\nBIT - SERIAL NEURAL  NETWORKS \n\nAlan...
1     1 \n\nCONNECTIVITY VERSUS ENTROPY \n\nYaser  S...
2     278 \n\nTHE HOPFIELD MODEL WITH MUL TI-LEVEL N...
3     442 \n\nAlan  Lapedes \nRobert  Farber \n\nThe...
4     740 \n\nSPATIAL  ORGANIZATION  OF  NEURAL  NEn...
...                                                 ...
9675  Discrete Object Generation\n\nwith Reversible ...
9676  Adaptively Aligned Image Captioning via\n\nAda...
9677  Fully Dynamic Consistent Facility Location\n\n...
9678  Efficient Rematerialization for Deep Networks\n...
9679  Flow-based Image-to-Image Translation\n\nwith ...

[9680 rows x 5 columns]>
```

```
authors_data.info
```

```
<bound method DataFrame.info of        source_id first_name    last_name
0             27       Alan       Murray                    NaN
1             27    Anthony        Smith                    NaN
2             27        Zoe       Butler                    NaN
3             63      Yaser  Abu-Mostafa                    NaN
4             60    Michael     Fleisher                    NaN
...          ...        ...          ...                    ...
30232       8693     Joshua         Wang                 Google
30233       2302       Ruho        Kondo  Toyota Central R&D Labs., Inc.
30234       2302    Keisuke       Kawano   Toyota Central R&D Labs., Inc
30235       2302    Satoshi        Koide        Toyota Central R&D Labs.
30236       2302     Takuro      Kutsuna   Toyota Central R&D Labs. Inc.

[30237 rows x 4 columns]>
```

- Print data:

```
authors_data
```

	source_id	first_name	last_name	institution
0	27	Alan	Murray	NaN
1	27	Anthony	Smith	NaN
2	27	Zoe	Butler	NaN
3	63	Yaser	Abu-Mostafa	NaN
4	60	Michael	Fleisher	NaN
...	...	...	...	...
30232	8693	Joshua	Wang	Google
30233	2302	Ruho	Kondo	Toyota Central R&D Labs., Inc.
30234	2302	Keisuke	Kawano	Toyota Central R&D Labs., Inc
30235	2302	Satoshi	Koide	Toyota Central R&D Labs.
30236	2302	Takuro	Kutsuna	Toyota Central R&D Labs. Inc.

```
papers_data
```

	source_id	year	title	abstract	full_text
0	27	1987	Bit-Serial Neural Networks	NaN	573 \n\nBIT - SERIAL NEURAL NETWORKS \n\nAlan...
1	63	1987	Connectivity Versus Entropy	NaN	1 \n\nCONNECTIVITY VERSUS ENTROPY \n\nYaser S...
2	60	1987	The Hopfield Model with Multi-Level Neurons	NaN	278 \n\nTHE HOPFIELD MODEL WITH MUL TI-LEVEL N...
3	59	1987	How Neural Nets Work	NaN	442 \n\nAlan Lapedes \nRobert Farber \n\nThe...
4	69	1987	Spatial Organization of Neural Networks: A Pro...	NaN	740 \n\nSPATIAL ORGANIZATION OF NEURAL NEn...
...	...	...	...	...	...
9675	5452	2019	Discrete Object Generation with Reversible Ind...	The success of generative modeling in continuo...	Discrete Object Generation\n\nwith Reversible ...
9676	4799	2019	Adaptively Aligned Image Captioning via Adapti...	Recent neural models for image captioning usua...	Adaptively Aligned Image Captioning via\n\nAda...
9677	1827	2019	Fully Dynamic Consistent Facility Location	We consider classic clustering problems in ful...	Fully Dynamic Consistent Facility Location\n\n...
9678	8693	2019	Efficient Rematerialization for Deep Networks	When training complex neural networks, memory ...	Efficient Rematerialization for Deep Networks\n...
9679	2302	2019	Flow-based Image-to-Image Translation with Fea...	Learning non-deterministic dynamics and intrin...	Flow-based Image-to-Image Translation\n\nwith ...

9680 rows × 5 columns

```
authors_data.dtypes
```

```
source_id       int64
first_name     object
last_name      object
institution    object
dtype: object
```

```
papers_data.dtypes
```

```
source_id     int64
year          int64
title        object
abstract     object
full_text    object
dtype: object
```

```
authors_data.isna().sum()
```

```
source_id          0
first_name         1
last_name          3
institution    12934
dtype: int64
```

```
papers_data.isna().sum()
```

```
source_id       0
year            0
title           0
abstract     3319
full_text       3
dtype: int64
```

```
authors_data.describe()
```

	source_id
count	30237.000000
mean	2135.749942
std	1952.749005
min	1.000000
25%	759.000000
50%	1526.000000
75%	2798.000000
max	9406.000000

```
papers_data.describe()
```

	source_id	year
count	9680.000000	9680.000000
mean	1963.827479	2009.498760
std	1825.720545	9.233312
min	1.000000	1987.000000
25%	717.000000	2003.000000
50%	1403.000000	2013.000000
75%	2579.000000	2018.000000
max	9406.000000	2019.000000

- Clean data :

```
# replacing na values
authors_data["institution"].fillna("No institution", inplace = True)
authors_data["first_name"].fillna("some one", inplace = True)
authors_data["last_name"].fillna("some one", inplace = True)
```

```
authors_data.isna().sum()
```

```
source_id      0
first_name     0
last_name      0
institution    0
dtype: int64
```

```
# replacing na values
papers_data['full_text'].fillna("no full_text ", inplace = True)
papers_data['abstract'].fillna("no abstract ", inplace = True)
```

```
papers_data.isna().sum()
```

```
source_id    0
year         0
title        0
abstract     0
full_text    0
dtype: int64
```

- Visualization

```
#Q1 \ find all paper write in 2019 ?
paper_title_2019= papers_data.loc[papers_data['year'] == 2019]
paper_title_2019
```

	source_id	year	title	abstract	full_text
8252	8547	2019	Compositional Plan Vectors	Autonomous agents situated in real-world envir...	Plan Arithmetic: Compositional Plan Vectors fo...
8253	610	2019	Learning to Propagate for Graph Meta-Learning	Meta-learning extracts the common knowledge fr...	Learning to Propagate for Graph Meta-Learning\...
8254	1164	2019	XNAS: Neural Architecture Search with Expert A...	This paper introduces a novel optimization met...	XNAS: Neural Architecture Search\n\nwith Exper...
8255	7845	2019	Multi-resolution Multi-task Gaussian Processes	We consider evidence integration from potentia...	Multi-resolution Multi-task Gaussian Processes...
8256	348	2019	Deep Equilibrium Models	We present a new approach to modeling sequenti...	Deep Equilibrium Models\n\nShaojie Bai\n\nJ. Z...
...	...	...	...	...	...
9675	5452	2019	Discrete Object Generation with Reversible Ind...	The success of generative modeling in continuo...	Discrete Object Generation\n\nwith Reversible ...
9676	4799	2019	Adaptively Aligned Image Captioning via Adapti...	Recent neural models for image captioning usua...	Adaptively Aligned Image Captioning via\n\nAda...
9677	1827	2019	Fully Dynamic Consistent Facility Location	We consider classic clustering problems in ful...	Fully Dynamic Consistent Facility Location\n\n...
9678	8693	2019	Efficient Rematerialization for Deep Networks	When training complex neural networks, memory ...	Efficient Rematerialization for Deep Networks\n...
9679	2302	2019	Flow-based Image-to-Image Translation with Fea...	Learning non-deterministic dynamics and intrin...	Flow-based Image-to-Image Translation\n\nwith ...

1428 rows × 5 columns

```
# Q2 \ find the most famous authors

n = 1
most_famous_authors= authors_data['first_name'].value_counts()[:n].index.tolist()
print('The most famous author Write about NIPS is :',most_famous_authors)
```

```
The most famous author Write about NIPS is : ['David']
```

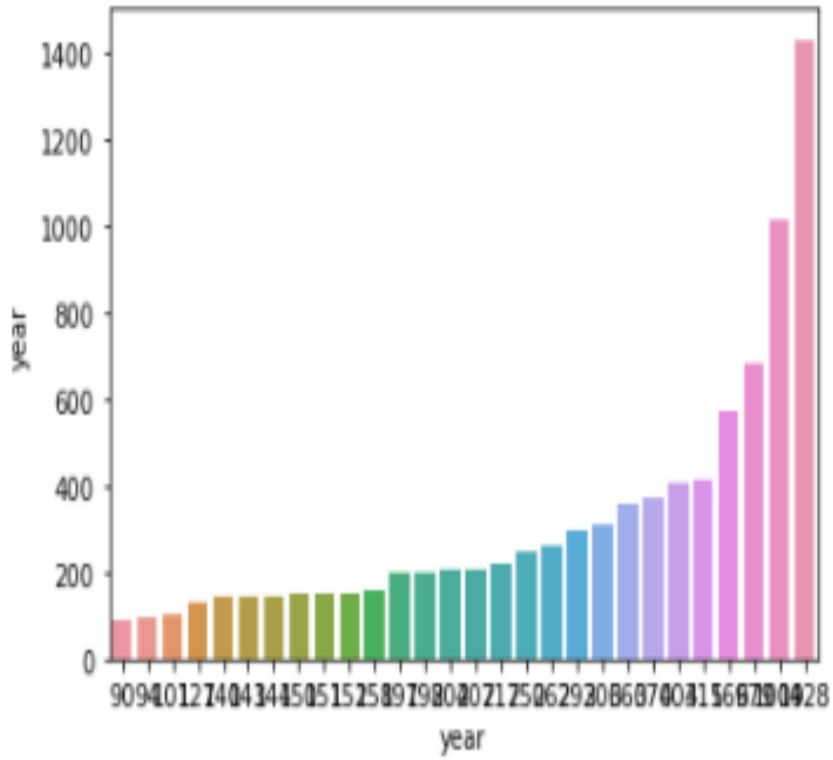
```
#Generating WordClouds in Python
n = 8
most_famous_authors= authors_data['first_name'].value_counts()[:n].index.tolist()
str1 = '  '.join(str(e) for e in most_famous_authors)

# Start with one review:
text = str1

# Create and generate a word cloud image:
wordcloud = WordCloud().generate(text)

# Display the generated image:
wordcloud = WordCloud(max_font_size=50, max_words=100, background_color="white").generate(text)
plt.figure()
plt.imshow(wordcloud, interpolation='bilinear')
plt.axis("off")
plt.show()
```

```
# Q3 \ find how many paper write in every year
num_paper_inyear= papers_data['year'].value_counts()
newdata= pd.DataFrame(num_paper_inyear)
newdata.columns.values[0] ='total'
newdata
```

	total
2019	1428
2018	1009
2017	679
2016	569
2014	411
2015	403
2012	370
2013	360
2011	306
2010	292
2009	262
2008	250
2007	217
2005	207
2004	207

2004	207
2002	207
2006	204
2003	198
2001	197
1993	158
1995	152
2000	152
1996	152
1998	151
1997	150
1999	150
1991	144
1990	143
1994	140
1992	127
1989	101
1988	94
1987	90

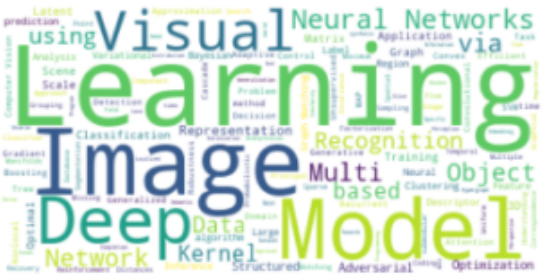
```
# Q3 \ find how many paper write in every year
num_paper_inyear= papers_data['year'].value_counts()
sns.barplot(x =num_paper_inyear,y=num_paper_inyear,data=num_paper_inyear);
```



```python
#Q4/ filtering the rows where paper abstract is about computer vision
df = papers_data[papers_data['abstract'].str.contains('computer vision')]
print(df)
papers_data['abstract'].str.contains(" computer vision").sum()
```

```
      source_id  year                                              title  \
3186        874  2007  The discriminant center-surround hypothesis fo...
3326        102  2007                Spatial Latent Dirichlet Allocation
3345        833  2008              Grouping Contours Via a Related Image
3351         60  2008  Cascaded Classification Models: Combining Mode...
3437        431  2008  Offline Handwriting Recognition with Multidime...
...         ...   ...                                                ...
9474       5717  2019  Unsupervised Learning of Object Keypoints for ...
9562       4676  2019       Graph Structured Prediction Energy Networks
9605       8988  2019  Neural Taskonomy: Inferring the Similarity of ...
9622       8790  2019  Certifying Geometric Robustness of Neural Netw...
9651       7255  2019  Online Convex Matrix Factorization with Repres...

                                               abstract  \
3186  The classical hypothesis, that bottom-up salie...
3326  In recent years, the language model Latent Dir...
3345  Contours have been established in the biologic...
3351  One of the original goals of computer vision w...
3437  Offline handwriting recognition---the transcri...
...                                                 ...
9474  The study of object representations in compute...
9562  For joint inference over multiple variables, a...
9605  Convolutional neural networks (CNNs) trained f...
9622  The use of neural networks in safety-critical ...
9622  The use of neural networks in safety-critical ...
9651  Matrix factorization (MF) is a versatile learn...

                                              full_text
3186  The discriminant center-surround hypothesis fo...
3326  Spatial Latent Dirichlet Allocation\n\nXiaogan...
3345  Grouping Contours Via a Related Image\n\nPrave...
3351  Cascaded Classification Models:\n\nCombining Mo...
3437  Offline Handwriting Recognition with\n\nMultidi...
...                                                 ...
9474  Unsupervised Learning of Object Keypoints\n\nf...
9562  Graph Structured Prediction Energy Networks\n\...
9605  Neural Taskonomy: Inferring the Similarity of\...
9622  Certifying Geometric Robustness of Neural Netw...
9651  Online Convex Matrix Factorization with\n\nRep...

[136 rows x 5 columns]

136
```

```
df = papers_data[papers_data['abstract'].str.contains('computer vision')]
#print(df.title)
tolist=df.title.tolist()
#type(df)
str1 = '  '.join(str(e) for e in tolist)

# Start with one review:
text = str1

# Create and generate a word cloud image:
wordcloud = WordCloud().generate(text)

# Display the generated image:
wordcloud = WordCloud(max_font_size=1000, max_words=600, background_color="white").generate(text)
plt.figure()
plt.imshow(wordcloud, interpolation='bilinear')
plt.axis("off")
plt.show()
```



By : Arwa Alolyani