

# En Sevdiğim Defterim 'Jupyter':)

*Gönüllü Aycı*

ÖYLG

12 Mayıs 2019

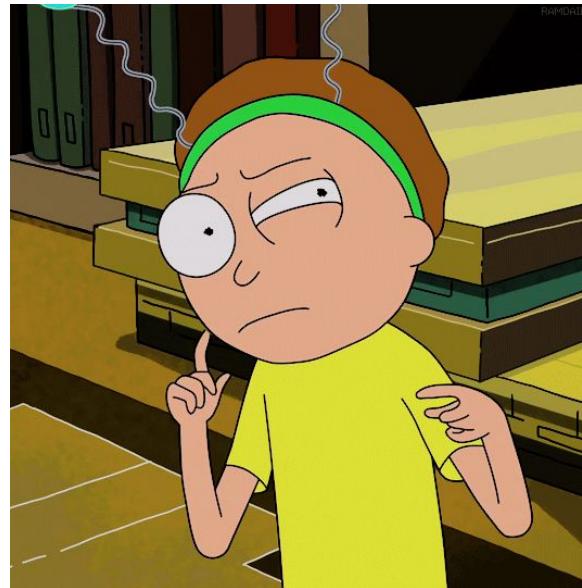




# Agenda

- ❖ What is Jupyter Notebook?
- ❖ How Jupyter Notebook works?
- ❖ What is Word2Vec?
  
- ❖ Twitter API example
- ❖ Pandas Dataframe example

# What is Jupyter Notebook?



# Jupyter Notebook



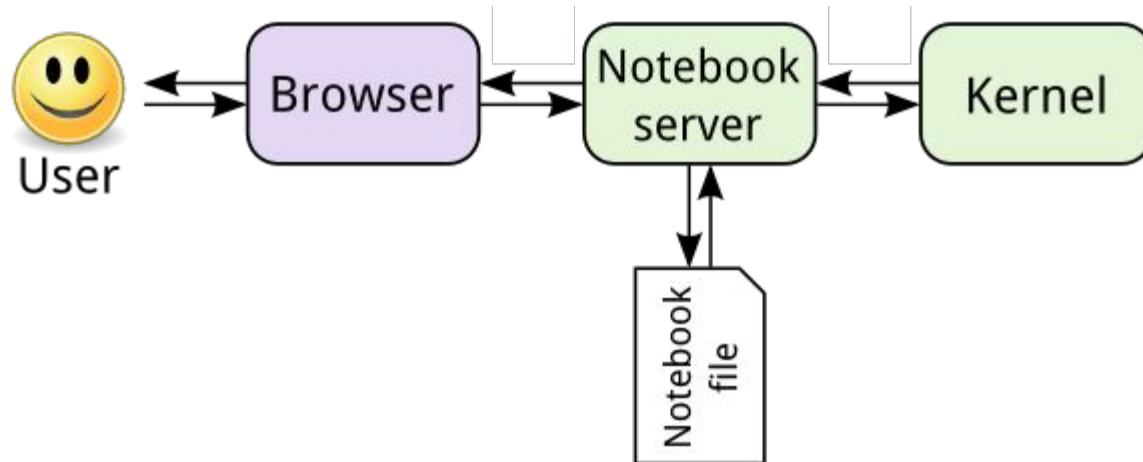
Fernando Pérez  
IPython → Jupyter Notebook  
2001 → 2014

- ❖ Interactive and shareable notebooks
  - Live code
  - Equations
  - Visualizations
  - Narrative text

# Jupyter Notebook



# How Jupyter Notebook works?



# DEMO: Jupyter Notebook



A screenshot of a web browser window titled "oylg2019" showing the Jupyter Notebook interface. The browser tabs include "oylg2019 - Google Slides", "oylg2019/", and "oylg2019". The address bar shows "localhost:8889/tree/oylg2019". The Jupyter logo and title "jupyter" are visible at the top. The interface includes navigation buttons, a toolbar with icons for file operations, and a menu bar with "Quit" and "Logout". Below the toolbar, there are tabs for "Files", "Running", and "Clusters", with "Files" selected. A message "Select items to perform actions on them." is displayed above the file list. The file list shows the contents of the directory "oylg2019", including "oylg2019.ipynb" (Running, a minute ago, 54.1 kB), "data.csv" (an hour ago, 1.4 MB), "fig\_mikolov3.jpg" (a month ago, 8.81 kB), and "trmodel" (a month ago, 664 MB). There are checkboxes next to each item, and columns for Name, Last Modified, and File size.

	Name	Last Modified	File size
<input type="checkbox"/>	..	seconds ago	
<input type="checkbox"/>	oylg2019.ipynb	Running a minute ago	54.1 kB
<input type="checkbox"/>	data.csv	an hour ago	1.4 MB
<input type="checkbox"/>	fig_mikolov3.jpg	a month ago	8.81 kB
<input type="checkbox"/>	trmodel	a month ago	664 MB

The screenshot shows a Jupyter Notebook interface running in a browser window titled 'oylg2019'. The notebook has a single open cell containing HTML code. Below the cell, several terminal commands are run, showing the current date, directory, and file operations.

```
1 <html>
2 <body>
3
4 <h1 style="color:DodgerBlue;"> Özgür Yazılım & Linux Günleri 2019 - En Sevdiğim defterim 'Jupyter' :D 🌱 └─</h1>
5
6 </body>
7 </html>
```

In [1]:

```
1 !date
```

Sun May 12 09:05:24 +03 2019

In [2]:

```
1 !pwd
```

/Users/brain/Documents/events/oylg2019

In [3]:

```
1 %mkdir deneme
2 %cd deneme
```

/Users/brain/Documents/events/oylg2019/deneme

In [4]:

```
1 !touch dosyal
2 !ls
```

dosyal

In [5]:

```
1 %cd ..
```

/Users/brain/Documents/events/oylg2019

The screenshot shows a Jupyter Notebook interface running on a Mac OS X desktop. The browser tabs are for 'oylg2019 - Google Slides', 'oylg2019/oylg2019.ipynb', and 'oylg2019'. The notebook title is 'jupyter oylg2019 Last Checkpoint: 13 hours ago (unsaved changes)'. The toolbar includes File, Edit, View, Insert, Cell, Kernel, Widgets, Help, Trusted, Python 3, and various icons for file operations like Open, Save, and Run.

The main content area displays a series of terminal commands run in cells:

- In [1]:  
1 !date  
Sun May 12 09:05:24 +03 2019
- In [2]:  
1 !pwd  
/Users/brain/Documents/events/oylg2019
- In [3]:  
1 %mkdir deneme  
2 %cd deneme  
/Users/brain/Documents/events/oylg2019/deneme
- In [4]:  
1 !touch dosyal  
2 !ls  
dosyal
- In [5]:  
1 %cd ..  
/Users/brain/Documents/events/oylg2019

oylg2019 - Google Slides x | oylg2019 x oylg2019 x +

localhost:8889/notebooks/oylg2019/oylg2019.ipynb

jupyter oylg2019 Last Checkpoint: 13 hours ago (unsaved changes)

Logout

File Edit View Insert Cell Kernel Widgets Help Trusted Python 3

In [6]: 1 !rm -r deneme/

In [7]: 1 %%writefile hello.py  
2 print("Hello everyone!")

Writing hello.py

In [8]: 1 %run hello.py

Hello everyone!

In [9]: 1 import pandas as pd

In [10]: 1 dataframe = pd.read\_csv('data.csv')  
2 dataframe

Out[10]:

	photoid	userid	normalizedpublic	url
0	4315936367	9adcf7e77f3c3ad6f53ac75a73de1bd2	1	https://farm5.staticflickr.com/4058/431593636... vector:0.9977343;illustration:0.9955312;re
1	4379045889	100b8cad7cf2a56f6df78f171f97a1ec	0	https://farm3.staticflickr.com/2791/437904588... performance:0.9972302;music:0.99513304;con
2	4373347844	134011f009df3c8f47cb522eb4fd28ae	1	https://farm5.staticflickr.com/4041/437334784... bird:0.99797726;wildlife:0.992786;no perso
3	4380747098	178572c23901d8548021c712dccabdb	0	https://farm5.staticflickr.com/4006/438074709... people:0.99083316;two:0.9721033;woman:0.971
4	4317062772	6f47e4d11a73cd41bc0311ac5ed4b4759	1	https://farm3.staticflickr.com/2706/431706277... no person:0.9957253;sea seal:0.98808414;wa
5	4328277168	2196b3ae4e772dc40dce82e1c54c2d7	1	https://farm5.staticflickr.com/4005/432827716... mammal:0.9983638;wildlife:0.99530536;no per
6	4335939833	f288a8ac2358d7c838eb879ba4d8d9f4	1	https://farm3.staticflickr.com/2717/433593983... tree:0.9986296;wood:0.9943332;winter:0.9931
7	4380898466	100b8cad7cf2a56f6df78f171f97a1ec	1	https://farm5.staticflickr.com/4042/438089846... computer:0.9951811;technology:0.9893943;ch

The screenshot shows a web browser window with three tabs open:

- oylg2019 - Google Slides
- oylg2019/
- oylg2019

The current tab is "oylg2019" which displays a Jupyter Notebook titled "oylg2019". The notebook interface includes:

- A toolbar with File, Edit, View, Insert, Cell, Kernel, Widgets, Help, Trusted, and Python 3 buttons.
- A toolbar with icons for file operations like New, Open, Save, Run, Cell Type, and Cell Mode.
- A code cell containing the following text:

```
1 # <center> LaTeX <center>
2
3 ## **Euclidean Distance**
```
- An output cell containing the following text:

```
1 $ \sqrt{ \sum_{i=1}^n (x_i - y_i)^2 } $
```
- An output cell containing the following LaTeX code:

```
1 \begin{equation}\sqrt{\sum_{i=1}^n (x_i - y_i)^2}\end{equation}
```

A screenshot of a web browser window displaying a Jupyter Notebook at `localhost:8889/notebooks/oylg2019/oylg2019.ipynb`. The browser has three tabs open: "oylg2019 - Google Slides", "oylg2019/", and the current notebook "oylg2019". The notebook interface includes a header with a Python logo, "Logout", "Trusted", and "Python 3" buttons. Below the header is a toolbar with various icons for file operations, cell types, and execution. The main content area contains a section titled "LaTeX" with a blue "¶" symbol, followed by the title "Euclidean Distance" and its mathematical definition.

## LaTeX

¶

### Euclidean Distance

$$\sqrt{\sum_{i=1}^n (x_i - y_i)^2}$$
$$\sqrt{\sum_{i=1}^n (x_i - y_i)^2}$$



Word2Vec

**Vector  
Representations  
of a Word**

**Word Embedding**

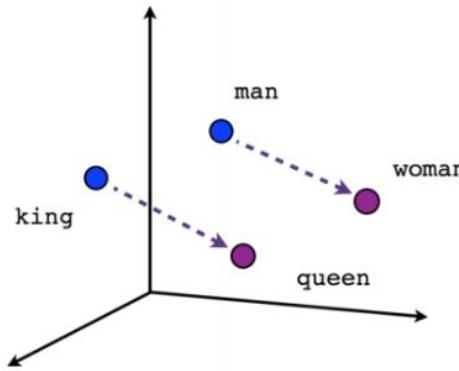


Tomas Mikolov  
Czech Computer Scientist  
2013 @ Google

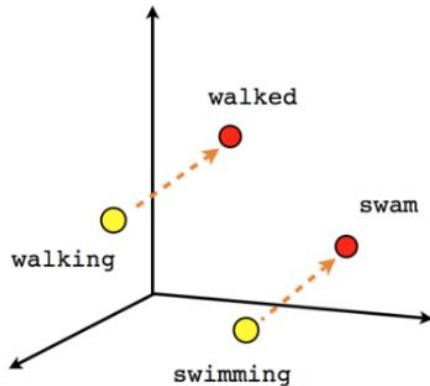
**One of the most  
popular technique to  
learn word  
embeddings using  
two-layer neural  
networks**

**Word2Vec**

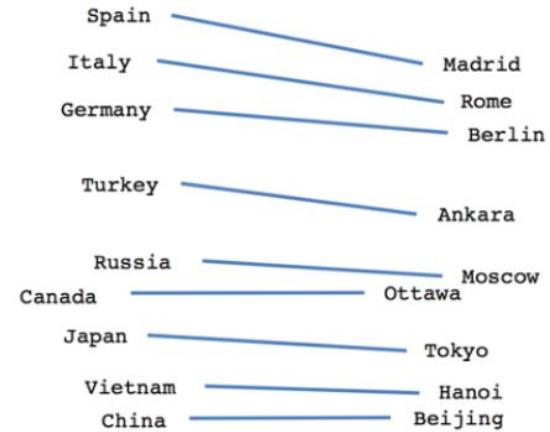
# Word2Vec



Male-Female

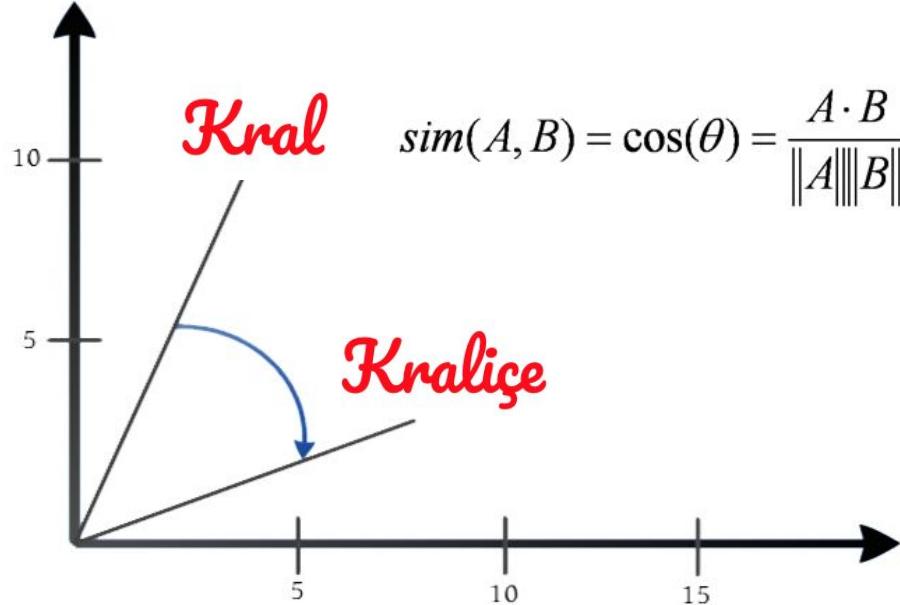


Verb tense



Country-Capital

# Cosine Similarity



# DEMO: Word2Vec



oylg2019 - Google Slides x | oylg2019/ x | oylg2019 x +

localhost:8889/notebooks/oylg2019/oylg2019.ipynb

Jupyter oylg2019 Last Checkpoint: 13 hours ago (unsaved changes)

Logout

File Edit View Insert Cell Kernel Widgets Help Trusted Python 3

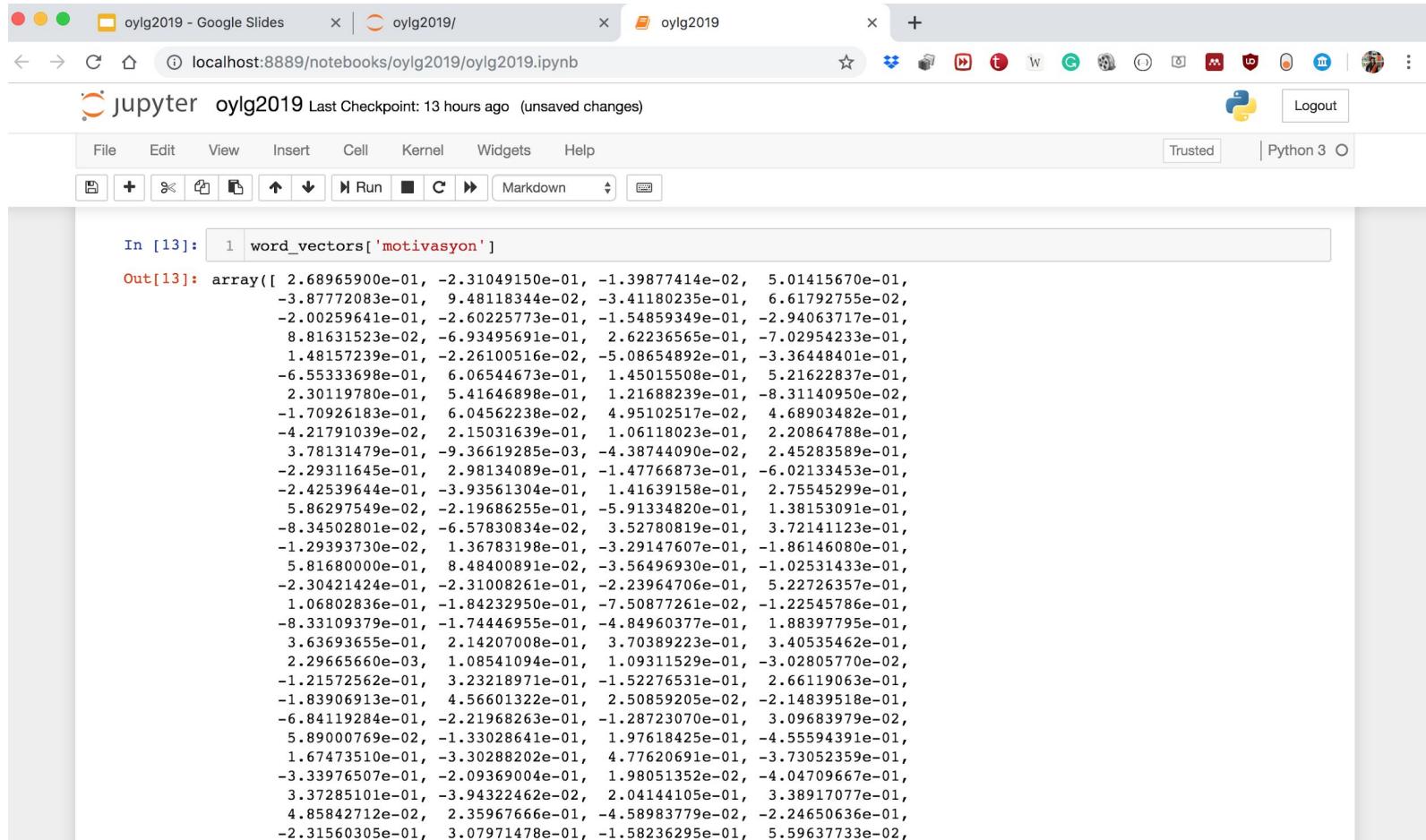
Word2Vec by Tomas Mikolov [1, 2]



```
In [11]: 1 import time  
2 from gensim.models import KeyedVectors
```

```
In [12]: 1 start = time.time()  
2  
3 word_vectors = KeyedVectors.load_word2vec_format('trmodel', binary=True)  
4  
5 print("total time: ", time.time() - start)  
6
```

total time: 9.24258804321289



oylg2019 - Google Slides x | oylg2019/ x | oylg2019 x +

localhost:8889/notebooks/oylg2019/oylg2019.ipynb

jupyter oylg2019 Last Checkpoint: 13 hours ago (autosaved)

Logout

File Edit View Insert Cell Kernel Widgets Help Trusted Python 3

In [6]: 1 word\_vectors.most\_similar('motivasyon')

Out[6]: [('bilişsel', 0.7156226634979248), ('farkındalığı', 0.715126097202301), ('yaratıcılık', 0.71113520860672), ('farkındalık', 0.7062397599220276), ('manipülasyon', 0.6943678855895996), ('esneklik', 0.685146689414978), ('sureçlerin', 0.6806826591491699), ('davranışsal', 0.6806298494338989), ('spekulatif', 0.6668925285339355), ('psikoterapi', 0.665993094442749)]

In [7]: 1 word\_vectors.most\_similar(positive=["kral", "kadın"], negative=["erkek"])

Out[7]: [('kraliçe', 0.508816123008728), ('kralı', 0.4718089699745178), ('kralın', 0.44512394070625305), ('kraliçesi', 0.4190150499343872), ('prenses', 0.40713071823120117), ('hükümdar', 0.4056030213832855), ('prens', 0.3978961706161499), ('kraliçenin', 0.39632880687713623), ('veliaht', 0.38219934701919556), ('tahtı', 0.3773398995399475)]



File Edit View Insert Cell Kernel Widgets Help

Trusted



Python 3



In [8]: 1 word\_vectors.most\_similar(positive=["geliyor", "gitmek"], negative=["gelmek"])

Out[8]: [('gidiyor', 0.592621922492981), ('gidiyorum', 0.5690363645553589), ('gidelim', 0.5637924671173096), ('geldim', 0.5413458347320557), ('bakıyor', 0.5373592972755432), ('gittim', 0.5343414545059204), ('gideceğim', 0.5251941680908203), ('geldik', 0.5050601363182068), ('geliyoruz', 0.5047824382781982), ('gider', 0.504159688949585)]

In [9]: 1 word\_vectors.doesnt\_match(["elma", "portakal", "çilek", "ev"])

Out[9]: 'ev'

In [10]: 1 from sklearn.metrics.pairwise import cosine\_similarity

In [11]: 1 cosine\_similarity(word\_vectors['kral'].reshape(1,400), word\_vectors['kralice'].reshape(1,400))

Out[11]: array([[0.6076636]], dtype=float32)

## References

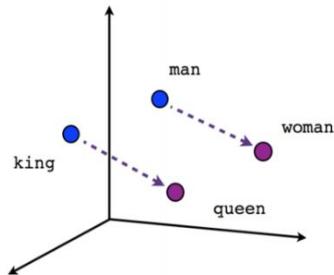
- [1] Mikolov, Tomas, et al. "Distributed representations of words and phrases and their compositionality." Advances in neural information processing systems. 2013.
- [2] Mikolov, Tomas, et al. "Efficient estimation of word representations in vector space." arXiv preprint arXiv:1301.3781 (2013).
- [3] <https://github.com/akoksal/Turkish-Word2Vec>
- [4] <https://nbviewer.jupyter.org/>
- [5] <http://tiny.cc/6gtj6y>

# References

- <https://nbviewer.jupyter.org/>
- <http://tiny.cc/jbvj6y>
- Mikolov, Tomas, et al. "Distributed representations of words and phrases and their compositionality." Advances in neural information processing systems. 2013.
- Mikolov, Tomas, et al. "Efficient estimation of word representations in vector space." arXiv preprint arXiv:1301.3781 (2013).
- **Word2Vec Turkish:** <https://github.com/akoksal/Turkish-Word2Vec>
- **Word2Vec English:** <https://nlp.stanford.edu/projects/glove/>
- <http://tiny.cc/6gtj6y>

# Thank You For Listening

## Any Questions?



aycignl



gonul\_ayci



aycignl