



Search Medium



▼

How to use Google Colab with GitHub via Google Drive

Mohammed Ismail P · [Follow](#)

Published in Analytics Vidhya

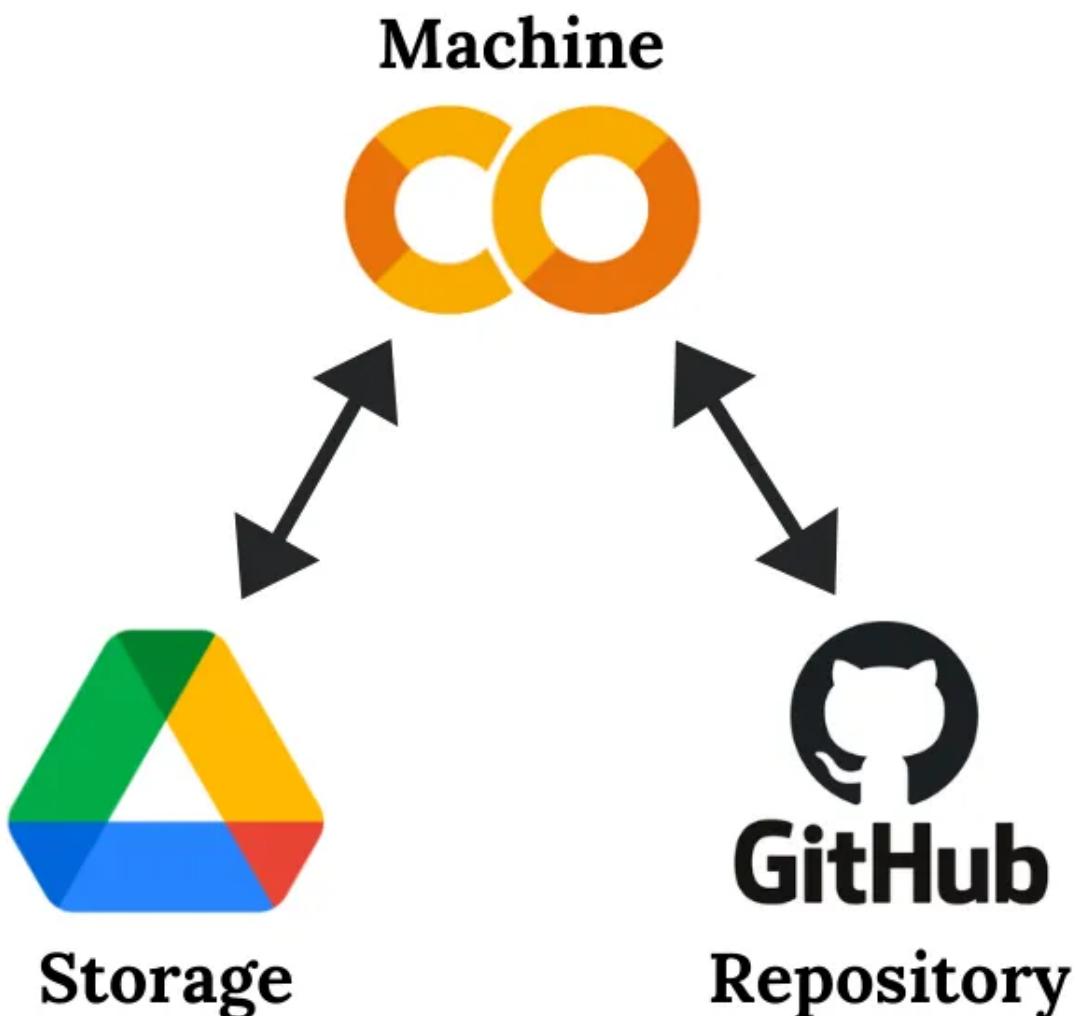
6 min read · Jan 30, 2021

[Listen](#)[Share](#)

In this tutorial, we will be discussing on how **Google Colab** can be used with **GitHub** for our Data science, Machine Learning projects and also use **Google Drive** as a cloud data storage. Let me introduce these products.

1. Colaboratory, or **Colab** for short, is a product from Google Research. It allows anybody to write and execute arbitrary python code through the browser, and is especially well suited to machine learning, data analysis and education.
2. **GitHub** is a code hosting platform for version control and Collaboration. It lets you and others work together on projects from anywhere. Thus allowing seamless collaboration without compromising the integrity of the original project.
3. **Google Drive** provides file storage and synchronization service, which allows users to store files on their servers, synchronize files across devices, and share files. It offers 15 GB of free storage to users.

Below picture depicts how these product interact with each other,



Step 1: Use colab notebook as a Shell Picture by Author — Interaction between GitHub, Colab and Drive

1. Visit [Google Colaboratory](#) website
2. Click on *New Notebook* button. A blank notebook is initialized and opened

Step 2: Mount Google Drive to Google Colab Notebook

- Run the below script to mount your Google Drive

```
from google.colab import drive  
drive.mount('/content/drive')
```

```
from google.colab import drive  
drive.mount('/content/drive')
```

Go to this URL in a browser:

Enter your authorization code:

Screenshot by Author — Authenticating Google account

- Click the link to authenticate user Google account
 - Select the respective Google Drive account on which you want to mount and click on sign in
 - Copy and Paste the authentication code into the input cell
 - Congrats! Your Google Drive is mounted.

```
[1] from google.colab import drive  
drive.mount('/content/drive')
```

Mounted at /content/drive

Screenshot by Author – Google Drive mounted

Step 3: Change present working directory

- Below shell command will set the present working directory to, /content/drive/MyDrive/Github

```
%cd /content/drive/MyDrive/Github/
```

Note: Your Google Drive's Home directory is at, /content/drive/MyDrive/

Step 4: Generate GitHub Access Token

Now its time to generate your GitHub token, that can be used to access the GitHub

API.

- Visit [GitHub](#) website and login to your account.
- Go to *Settings*, navigate to *Developer settings* and then click on *Personal access tokens*. Your page should look something like,

The screenshot shows the GitHub 'Personal access tokens' page. On the left, there's a sidebar with 'GitHub Apps', 'OAuth Apps', and 'Personal access tokens'. The 'Personal access tokens' option is selected and highlighted with a red border. The main area is titled 'Personal access tokens' and contains a sub-header 'Tokens you have generated that can be used to access the GitHub API.' Below this, two tokens are listed:

- Colab and Drive — repo**: Last used within the last week. A 'Delete' button is to the right.
- git: https://github.com/ on HP at 16-May-2020 23:19 — gist, repo, workflow**: Last used within the last 9 months. A 'Delete' button is to the right.

At the bottom, a note states: 'Personal access tokens function like ordinary OAuth access tokens. They can be used instead of a password for Git over HTTPS, or can be used to authenticate to the API over Basic Authentication.'

© 2021 GitHub, Inc. [Terms](#) [Privacy](#) [Security](#) [Status](#) [Docs](#) [Contact GitHub](#) [Pricing](#) [API](#) [Training](#) [Blog](#) [About](#)

Screenshot by Author — <https://github.com/settings/tokens>

- Click on *Generate new token* button on top right corner of the page.
- Click the *repo* checkbox under *Select scopes* section as shown,

[Settings](#) / Developer settings

[GitHub Apps](#)

[OAuth Apps](#)

[Personal access tokens](#)

New personal access token

Personal access tokens function like ordinary OAuth access tokens. They can be used instead of a password for Git over HTTPS, or can be used to [authenticate to the API over Basic Authentication](#).

Note

What's this token for?

Select scopes

Scopes define the access for personal tokens. [Read more about OAuth scopes](#).

<input checked="" type="checkbox"/> repo	Full control of private repositories
<input type="checkbox"/> repo:status	Access commit status
<input type="checkbox"/> repo_deployment	Access deployment status
<input type="checkbox"/> public_repo	Access public repositories
<input type="checkbox"/> repo:invite	Access repository invitations
<input type="checkbox"/> security_events	Read and write security events
<input type="checkbox"/> workflow	Update github action workflows
<input type="checkbox"/> write:packages	Upload packages to github package registry
<input type="checkbox"/> read:packages	Download packages from github package registry

Screenshot by Author — Generating new access token

Learn more about [scopes of access token](#).

- Now, click on *Generate token* button at the bottom of the page. Now, the page should look like,

[Settings](#) / Developer settings

[GitHub Apps](#)

[OAuth Apps](#)

[Personal access tokens](#)

Personal access tokens

[Generate new token](#) [Revoke all](#)

Tokens you have generated that can be used to access the [GitHub API](#).

Make sure to copy your new personal access token now. You won't be able to see it again!

<input checked="" type="checkbox"/> f10e3c5737e39d4f5e53b1d22a9834d81626f0b9	Copy	Delete
Colab and Drive — repo	Last used within the last week	Delete
git: https://github.com/ on HP at 16-May-2020 23:19 — gist, repo, workflow	Last used within the last 9 months	Delete

**Newly generated
access token**



Personal access tokens function like ordinary OAuth access tokens. They can be used instead of a password for Git over HTTPS, or can be used to [authenticate to the API over Basic Authentication](#).

Screenshot by Author — Generated access token

You have successfully generated access token for your GitHub account. We will be using this token to access GitHub API. **Note:** Do not share your access token to the public as I did, I'll be using for the purpose of this tutorial only.

Now, there arises two different scenario.

1. Create a **new git repository** from scratch
2. Clone an **existing git repository** from GitHub

Based on your requirements, follow the below steps,

Step 5.A: Create a new Git repository

Follow the below steps to create a new git repository from scratch directly in your Google Drive,

Step 5.A.1: Initialize new Git repository

- Initialize git using `git init <directory>` . In this tutorial, we will be using *titanic* repository.
- Change your working directory to the created repository.
- List the files and folder using `ls` command.

```
!git init titanic
```

```
Initialized empty Git repository in /content/gdrive/My Drive/Github/titanic/.git/
```

```
%cd titanic/
```

```
/content/gdrive/My Drive/Github/titanic
```

```
%ls -a
```

```
.git/
```

Screenshot by Author — git init command

Step 5.A.2: Working with Git repository

- It's time to add files and folders to your working directory.
- `git status` to view the state of the working directory and the staging area.
- `git add` to add changes in the working directory to the staging area.

Learn more about [How to download Kaggle datasets to Google Colab](#).

```
Added titanic notebook.ipynb notebook and titanic.csv file
```

```
!git status
```

```
On branch master
```

```
No commits yet
```

```
Untracked files:
```

```
(use "git add <file>..." to include in what will be committed)
```

```
    titanic notebook.ipynb  
    titanic.csv
```

```
nothing added to commit but untracked files present (use "git add" to track)
```

```
!git add .
```

```
!git status
```

```
On branch master
```

```
No commits yet
```

```
Changes to be committed:
```

```
(use "git rm --cached <file>..." to unstage)
```

```
    new file:   titanic notebook.ipynb  
    new file:   titanic.csv
```

Screenshot by Author — Adding files to working directory

- After adding files and folder as per your requirements, commit your work using

```
git commit -m "message" .
```

Step 5.A.3: Create a new repository on GitHub

Once you are satisfied with your work and want to save your commits to GitHub, follow the below steps,

- Visit [GitHub official](#) website and login with your account.
- Create a new repository in GitHub. **Note:** Do not initialize the repository with `README`, `.gitignore` or `license` file. This empty repository will await your code. Don't worry, you can create those files in Google colab (local machine) or after your first `git push` on GitHub.

Create a new repository

Owner *



Repository name *

Great repository names are short and memorable. Need inspiration? How about [fantastic-train?](#)

Description (optional)

Public

Anyone on the internet can see this repository. You choose who can commit.

Private

You choose who can see and commit to this repository.

Initialize this repository with:

Skip this step if you're importing an existing repository.

Add a README file

This is where you can write a long description for your project. [Learn more.](#)

Add .gitignore

Choose which files not to track from a list of templates. [Learn more.](#)

Choose a license

A license tells others what they can and can't do with your code. [Learn more.](#)

Screenshot by Author — Create a new repository on GitHub

Step 5.A.4: Upload your commits to GitHub

Before pushing your commits to your GitHub account, you need to configure your Google Drive git repository. To do this follow the steps,

- Create a set of variables from your GitHub account,

<https://github.com/<username>/<repository>>

1. `username` — Your GitHub username. In our case, its `MohammedIsmailP`.
 2. `repository` — Created repository. In our case, its `titanic`.
 3. `git_token` — Your personal access token (*Do not share to public*).
- Add `remote` to your git from the above variable as,

```
git remote add <remote-name> https://{}@github.com/{username}/{repository}.git
```

- Push your commits using `git push` command as,

```
git push -u <remote-name> <branch-name>
```

```
username = 'MohammedIsmailP'
git_token = 'f10e3c5737e39d4f5e53b1d22a9834d81626f0b9'
repository = 'titanic'

!git remote add origin https://{}@github.com/{username}/{repository}.git
!git remote -v

origin  https://f10e3c5737e39d4f5e53b1d22a9834d81626f0b9@github.com/MohammedIsmailP/titanic.git (fetch)
origin  https://f10e3c5737e39d4f5e53b1d22a9834d81626f0b9@github.com/MohammedIsmailP/titanic.git (push)

!git push -u origin master

Counting objects: 4, done.
Delta compression using up to 2 threads.
Compressing objects: 100% (4/4), done.
Writing objects: 100% (4/4), 22.12 KiB | 3.16 MiB/s, done.
Total 4 (delta 0), reused 0 (delta 0)
To https://github.com/MohammedIsmailP/titanic.git
 * [new branch]      master -> master
Branch 'master' set up to track remote branch 'master' from 'origin'.
```

Screenshot by Author — Pushing commits to GitHub

Congratulations!! You have successfully pushed your commits to GitHub. Let's verify it in GitHub.

MohammedIsmailP / titanic

Code Issues Pull requests Actions Projects Wiki Security Insights Settings

master 1 branch 0 tags Go to file Add file Code

MohammedIsmailP Added titanic notebook and titanic.csv file ff4a241 23 minutes ago 1 commit

titanic notebook.ipynb Added titanic notebook and titanic.csv file 23 minutes ago

titanic.csv Added titanic notebook and titanic.csv file 23 minutes ago

Help people interested in this repository understand your project by adding a README. Add a README

Screenshot by Author — git commits reflected in GitHub

Step 5.B: Clone an existing Git repository

Follow the below steps to clone an existing git repository from GitHub into your Google Drive,

- Go to your GitHub repository to clone the repository.
- Click on *Code* button and copy the url as shown,

MohammedIsmailP / titanic

Code Issues Pull requests Actions Projects Wiki Security Insights Settings

master 1 branch 0 tags Go to file Add file Code

MohammedIsmailP Added titanic notebook and titanic.csv file

titanic notebook.ipynb Added titanic notebook and titanic.csv file

titanic.csv Added titanic notebook and titanic.csv file

Clone
HTTPS SSH GitHub CLI
<https://github.com/MohammedIsmailP/titanic>

Use Git or checkout with SVN using the web URL.

Open with GitHub Desktop

Download ZIP

Help people interested in this repository understand your project by adding a README

Screenshot by Author — GitHub repository page

- You'll need your GitHub access token before cloning your GitHub repository.
Also extract set of variables from your GitHub account.

1. `username` — Your GitHub username. In our case, its `MohammedIsmailP`.
2. `repository` — Created repository. In our case, its `titanic`.
3. `git_token` — Your personal access token (*Do not share to public*).

```
git clone https://{git_token}@github.com/{username}/{repository}
```

```
username = 'MohammedIsmailP'  
repository = 'titanic'  
git_token = 'f10e3c5737e39d4f5e53b1d22a9834d81626f0b9'
```

```
!git clone https://{git_token}@github.com/{username}/{repository}
```

```
Cloning into 'titanic'...  
remote: Enumerating objects: 4, done.  
remote: Counting objects: 100% (4/4), done.  
remote: Compressing objects: 100% (4/4), done.  
remote: Total 4 (delta 0), reused 4 (delta 0), pack-reused 0  
Unpacking objects: 100% (4/4), done.
```

```
%cd {repository}
```

```
/content/gdrive/My Drive/Github/titanic
```

```
%ls -a
```

```
.git/ titanic.csv 'titanic notebook.ipynb'
```

Screenshot by Author — git clone with GitHub access token

Congratulations!! You have successfully cloned your GitHub repository. You can start working on your repository and save your works by committing them using `git commit -m "message"`.

Modified titanic notebook.ipynb notebook.

```
!git status
```

```
On branch master
Your branch is up to date with 'origin/master'.

Changes not staged for commit:
  (use "git add <file>..." to update what will be committed)
  (use "git checkout -- <file>..." to discard changes in working directory)

        modified:   titanic notebook.ipynb

no changes added to commit (use "git add" and/or "git commit -a")
```

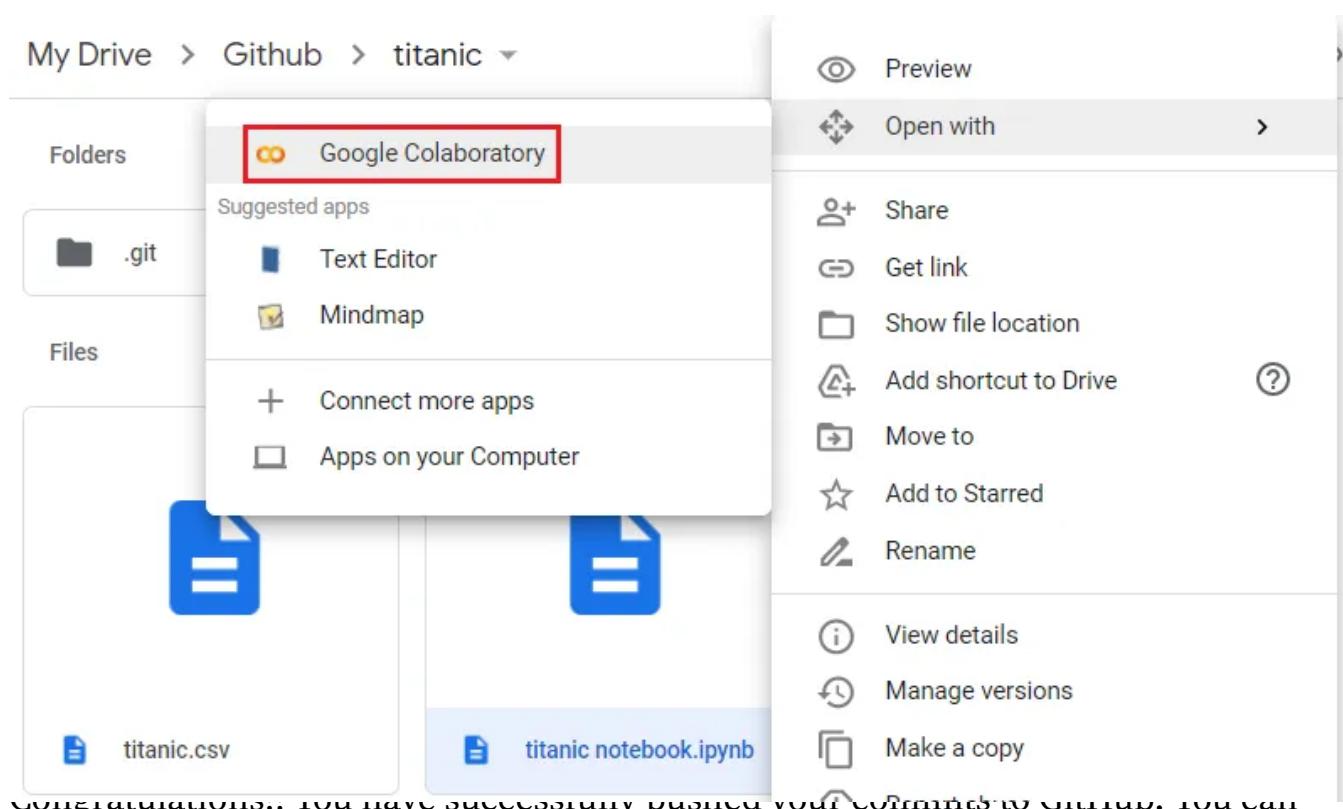
How to run notebook stored in Google Drive

```
!git add
```

To run your notebook stored in Google Drive, follow the steps.

- Open your Google Drive and login with your account.
- Locate your notebook in Google Drive and right click on the notebook to open actions menu, hover over *Open with*, and select *Google Colaboratory* as shown,

Screenshot by Author — Committing changes to Git



verify the screenshot by Author — Open notebook from Google Drive to Google Colaboratory

MohammedlsmailP / titanic

Code Issues Pull requests Actions Projects Wiki Security Insights Settings

master 1 branch 0 tags Go to file Add file Code

MohammedlsmailP Modified titanic notebook.ipynb	6ec5b4e 11 minutes ago	2 commits
titanic notebook.ipynb	Modified titanic notebook.ipynb	11 minutes ago
titanic.csv	Added titanic notebook and titanic.csv file	14 hours ago

Help people interested in this repository understand your project by adding a README. Add a README

That's it! Our tutorial Colab with GitHub via tutorial. But wait, we Don't worry, I'll cover this as well.

Screenshot by Author — Mount Google Drive and Change working directory

How to use Google along with this d in Google Drive'.

- Then execute your code.

That's it! Simple, isn't it? Hope, it helps. Happy Learning!!

Reference

1. [Scopes for OAuth Apps](#)
2. [How to download Kaggle datasets to Google Colab](#)
3. Learn more about Git and GitHub from [Git Handbook](#)

Machine Learning

Google Drive

Google Colab

Github

Version Control

[Follow](#)

Written by Mohammed Ismail P

29 Followers · Writer for Analytics Vidhya

Software Engineer at Zoho | Computer Science student at SRM Valliammai Engineering College | Data Analytics | Machine Learning | Deep Learning | Data Science

More from Mohammed Ismail P and Analytics Vidhya

```
datasets download -d heptapod/titanic --unzip
```

```
g titanic.zip to /content/gdrive/My Drive/Kaggle  
10.8k [00:00<?, ?B/s]  
:/10.8k [00:00<00:00, 1.56MB/s]
```

```
in train_and_test2.csv
```



Mohammed Ismail P in Analytics Vidhya

How to download Kaggle Datasets into Google Colab via Google Drive

... the easiest way to download datasets from Kaggle to Google Colab via Google Drive. Google Drive is used to store datasets for later use ...

5 min read · Jan 17, 2021



Kia Eisinga in Analytics Vidhya

How to create a Python library

Ever wanted to create a Python library, albeit for your team at work or for some open source project online? In this blog you will learn...

7 min read · Jan 26, 2020



1.97K



23





Harikrishnan N B in Analytics Vidhya

Confusion Matrix, Accuracy, Precision, Recall, F1 Score

Binary Classification Metric

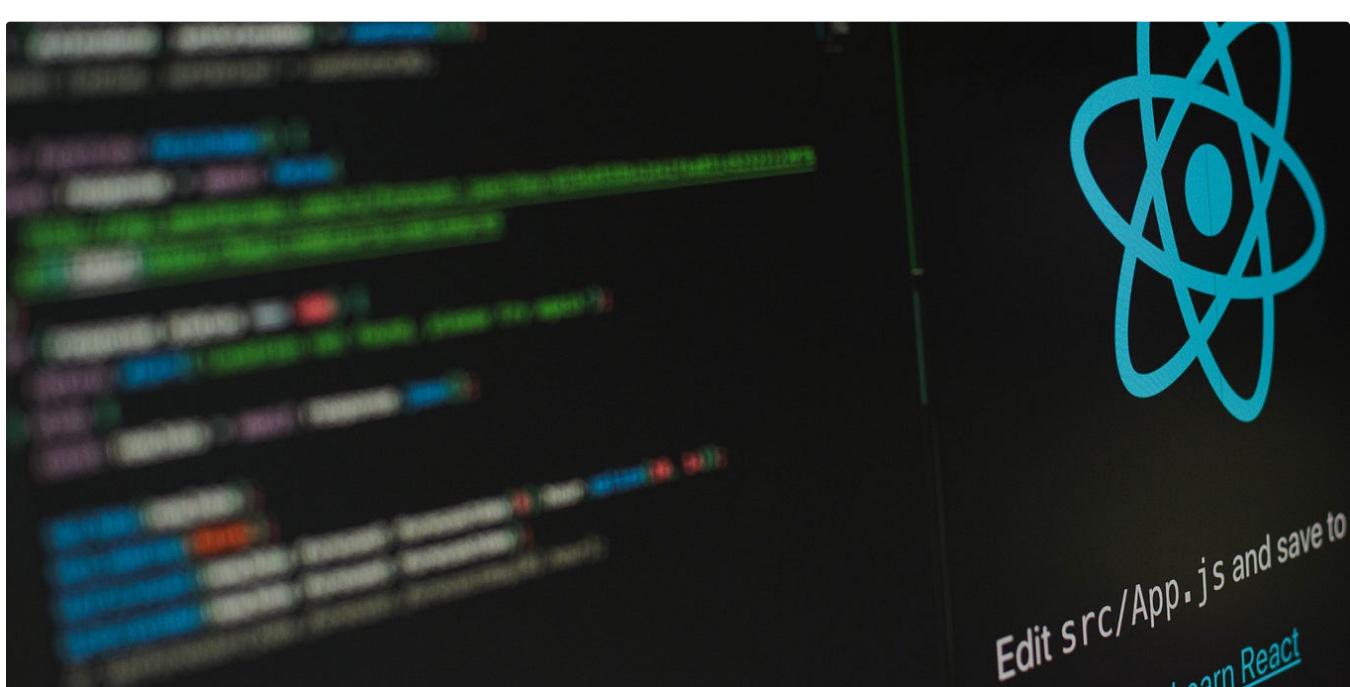
6 min read · Dec 10, 2019



649



6





Mohammed Ismail P

Create React App without `npx create-react-app`

As a React JS enthusiast, You may wonder how can you create a React App without the conventional `npx create-react-app` ...

8 min read · Sep 18, 2022



50



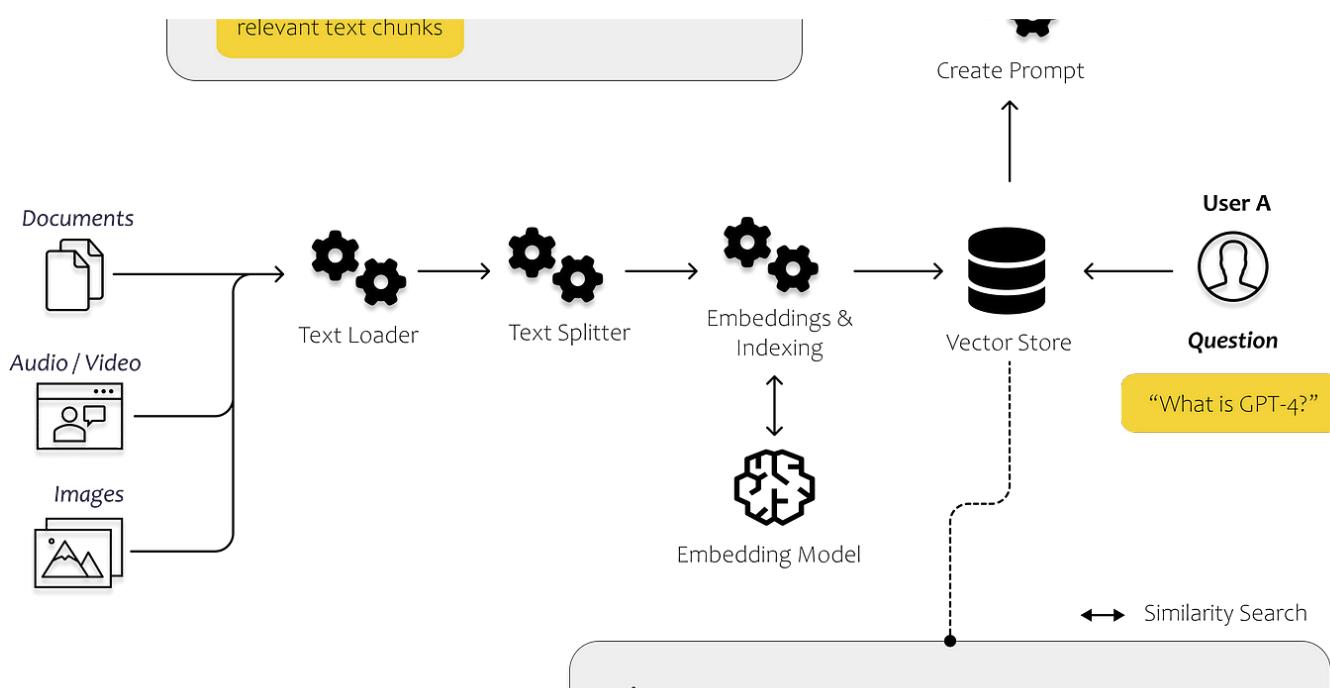
1



[See all from Mohammed Ismail P](#)

Recommended from Medium

[See all from Analytics Vidhya](#)



Dominik Polzer in Towards Data Science

All You Need to Know to Build Your First LLM App

A step-by-step tutorial to document loaders, embeddings, vector stores and prompt templates



· 26 min read · Jun 22

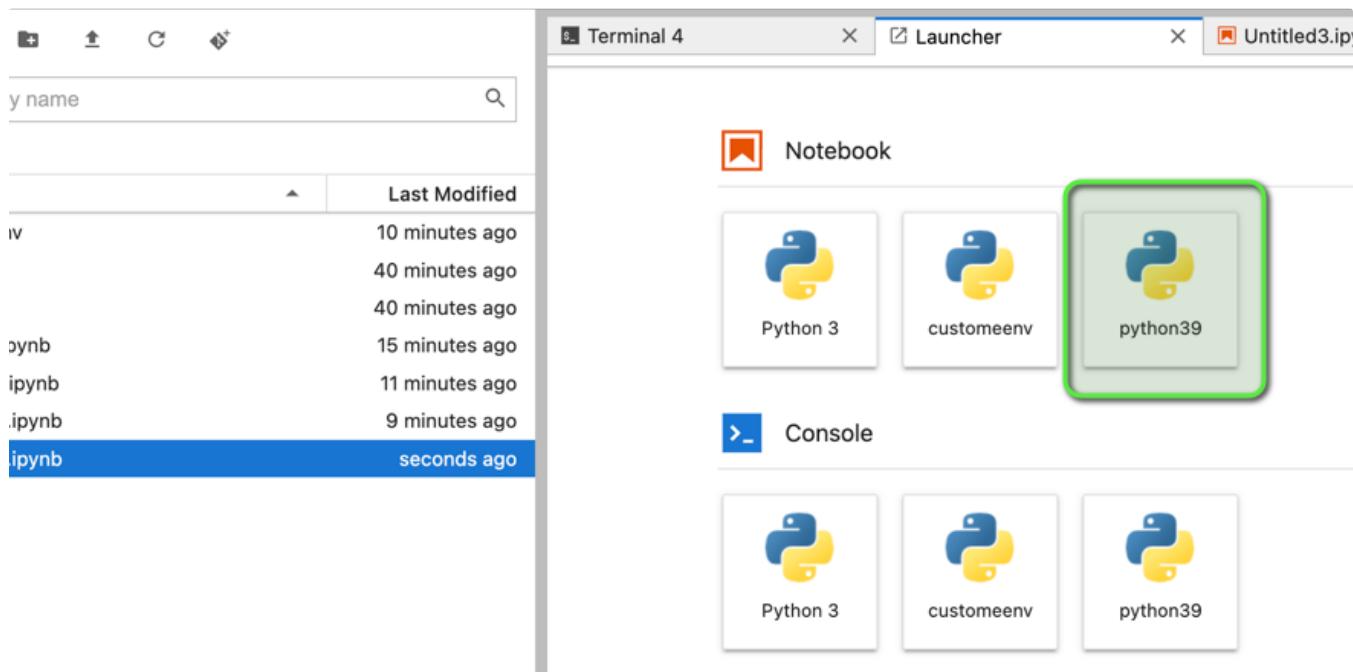


4.3K



40





Sascha Heyer in Google Cloud - Community

Upgrade Google Vertex AI Workbench Notebook Python Version

Google's Vertex AI Notebooks are currently running on Python 3.7. You might need a more recent Python version to use newer Python features.

◆ · 1 min read · Mar 1

80 4

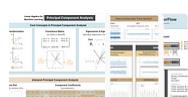


Lists



Predictive Modeling w/ Python

18 stories · 260 saves



Practical Guides to Machine Learning

10 stories · 270 saves



Natural Language Processing

495 stories · 129 saves



The New Chatbots: ChatGPT, Bard, and Beyond

13 stories · 82 saves



 Leonie Monigatti in Towards Data Science

Getting Started with LangChain: A Beginner's Guide to Building LLM-Powered Applications

A LangChain tutorial to build anything with large language models in Python

★ · 12 min read · Apr 25

 3.7K  28





Matt Chapman in Towards Data Science

The Portfolio that Got Me a Data Scientist Job

Spoiler alert: It was surprisingly easy (and free) to make

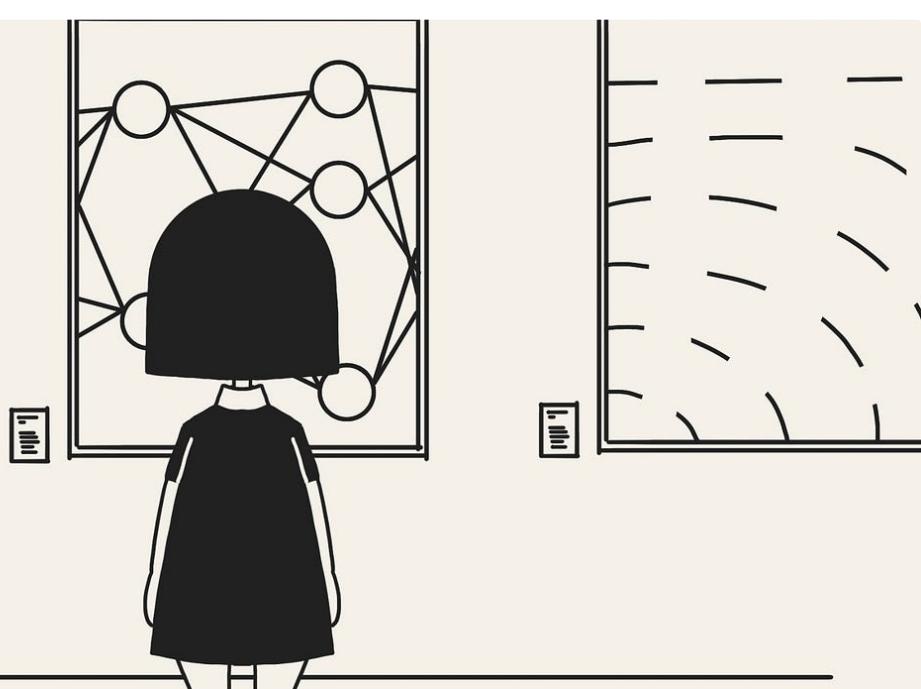
★ · 10 min read · Mar 24

4.2K

72



```
000100010001000100010001111  
0011001111000001111010000  
01111000101101110011110000/  
101001010001000100010001000  
110101010111011011110101010  
101011010101010111101110001  
0110101010001000001000100  
010002000001000100010001000  
100011110011101111000001111  
101000001111100010110111001  
1110000/10100101000100010001  
000100011010101011101101111  
010101010101101010101011110  
1110001011010101010001000000  
100010010002000001000100010  
001000100011110011101111000.
```



 Leonie Monigatti in Towards Data Science

10 Exciting Project Ideas Using Large Language Models (LLMs) for Your Portfolio

Learn how to build apps and showcase your skills with large language models (LLMs). Get started today!

★ · 11 min read · May 15

 1.98K 11 Timothy Mugayi in Better Programming

How To Build Your Own Custom ChatGPT With Custom Knowledge Base

Feed your ChatGPT bot with custom data sources

★ · 11 min read · Apr 7

 4.3K 99

See more recommendations

