How to run our Airbnb project

4/17/2023

These instructions inform on how to get up and running the following items in our GitHub repository https://github.com/clgong/airbnbapp

- Search and recommender App for Airbnb properties in Seattle, WA
- Jupyter notebooks that supported our research on the search and recommender app

Note: The steps provided are how to get things installed running on macOS Big Sur or later and briefly Windows is mentioned when there is an obvious difference. The focus is on plain vanilla Python and the standard Pip and Venv environments. For other specific OS's and python packages managers such as Conda, you'll need to do the equivalent steps, but they are not specifically covered here.

Prerequisites

- Access to github
- Python installed on your local machine, a recent version such as 3.8.2 or greater
    - pip (installed by default with python3 already)
    - venv (installed by default with python 3.3 or later already)
- Streamlit account (optional, if you want to deploy to Streamlit)
- AWS account (optional, if you want to deploy to Amazon Web Services)

Search and recommender App

Step 0. What to do if something doesn't work.

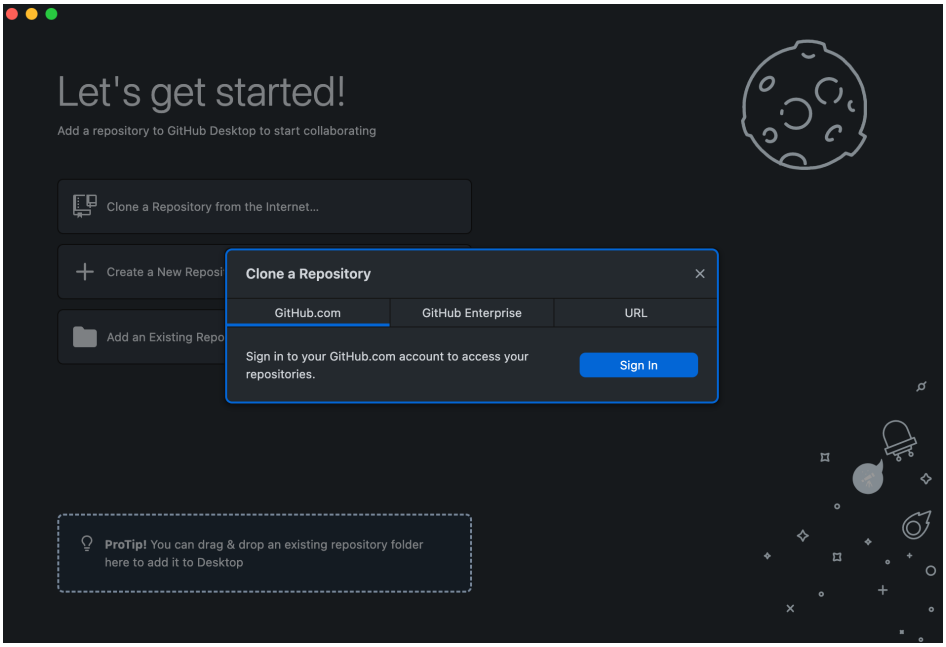
- The world of software and machine learning changes very quickly, so installation instructions can go out of date very fast. The **most** important step is to Google it, when something doesn't work.

Step 1. Check out the source code

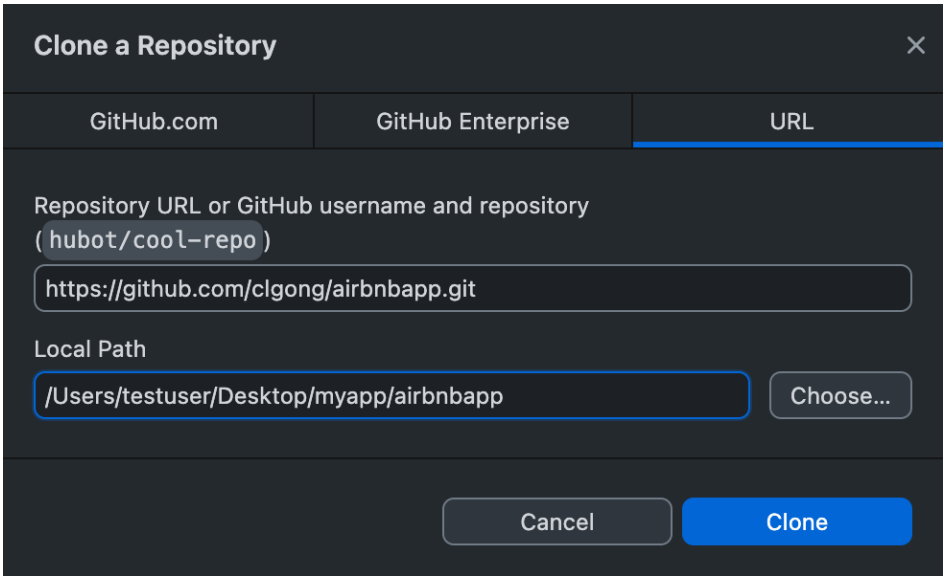
- Ensure you have access to github and can check out the project (located here https://github.com/clgong/airbnbapp) to your desktop folder "`myapp`".
    - **If you already know what you're doing in git** then you can ignore the step below. Just check out the code to a folder of your choice. We decided to use "`myapp`".
    - One way to do this is to use GitHub Desktop https://desktop.github.com/ …which is an app you can use on Windows, MacOS and Linux. It is a graphical user

interface so you don't have to remember all the git commands. That's what we used for this project.

- Download and install the GitHub Desktop app and run it.
- Click "Clone a repository from the internet…" or if you've already used the app before…the File menu "Clone Repository…" and select the URL tab



- Sign in with your GitHub account
- Use the UI to check out the project to a folder of your choice. (We used our Desktop and created "`myapp`" folder). Then check out the code to "`airbnbapp`".
- The project URL is "https://github.com/clgong/airbnbapp.git"

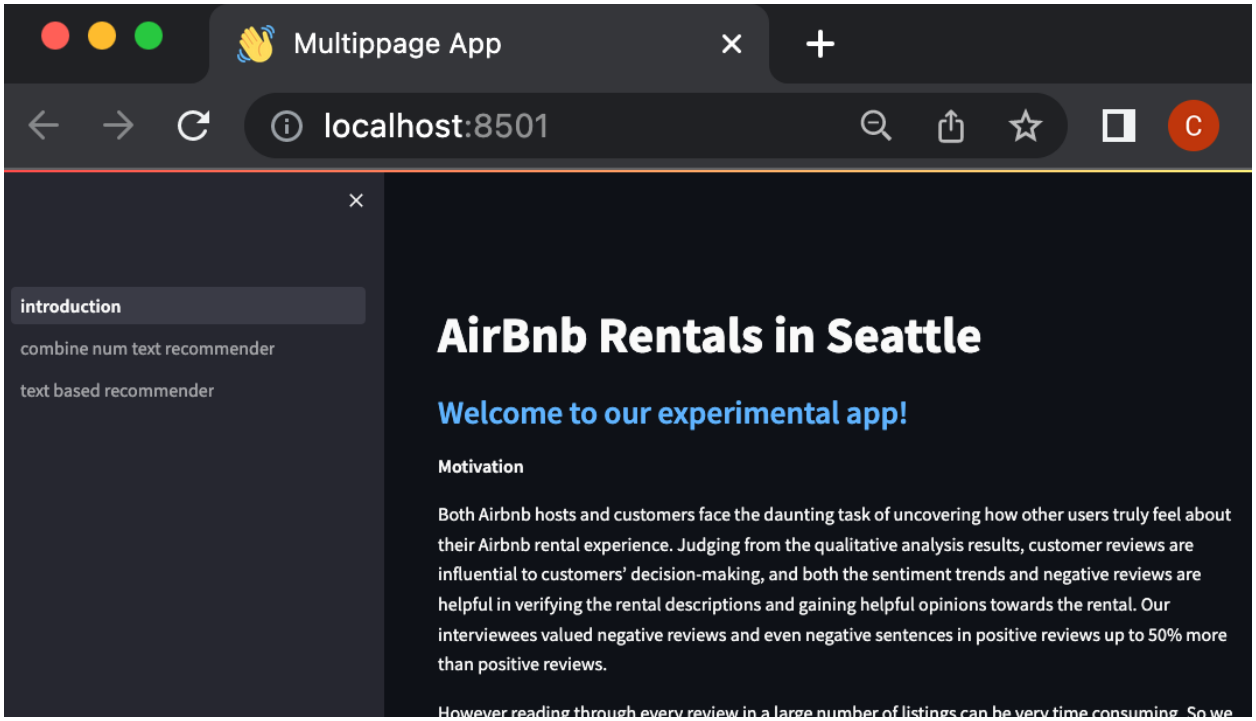


Step 2. Install the app's python library requirements (includes Streamlit) locally

- Set up a python virtual environment. You don't HAVE to do this but this is often done so that this particular project won't cause any library or dependency versioning problems with your other python projects on your computer. In our case we created it as "`vtest1`" but you can call it anything:
    - Open a terminal window.
        - Go to your source code directory. In our case it is "`Desktop/myapp/airbnbapp`"
        - Type this at the command prompt: `python3 -m venv vtest1`
        - Done!
        - NOTE: if you get an error "`No module named venv`" then you need to install venv first by doing this: `pip install venv`
- Activate the virtual environment in the terminal window. **NOTE**: replace `vtest1` with the name you chose:
    - In the same directory above, type this at the command prompt:
        - `source vtest1/bin/activate`
    - Notice the prompt says "`(vtest1)`" for example on our test computer "`(vtest1) TestUser-MacBook-Pro`"...this means you are in your virtual environment. Everything you install will now only go to this environment.
- Install python libraries needed for this project (includes Streamlit):
    - Type this at the command prompt: `pip install -r requirements.txt`
- NOTE: Once you are done working on your project for the day you can always exit the virtual environment by typing this: "`deactivate`" then later activate it again with the "`source`" command noted earlier.

Step 3. Run our airbnb app locally

- In your source code directory, type this: `streamlit run introduction.py`
- It will open a browser window to run your airbnb app
- NOTE: If you get an error with `seaborn` or `sklearn`, you'll need to pip install those by hand sometimes. This might be a bug the earlier step with `requirements.txt` should have already covered this (See text file for exact library versions used).
    - Quit streamlit with Control C at the keyboard in your terminal window. Then install these and run the app again (all terminal commands):
        - `pip install seaborn`
        - `pip install scikit-learn`
        - `streamlit run introductio.py`
- Once it's running you don't need to quit the browser every time you change code…just refresh the browser.
- Here is what the app should look like (http://localhost:8501/):



Step 4: Optional. Deploying your own app to the web hosted by Streamlit for free

- The previous steps showed how to check out the code and deploy the app locally. But if you want to deploy your app to the web, you should first do the steps below: (**Note**: It's beyond the scope of this document to instruct on how to fork the project, make your code changes, push to github.)

1. Fork the project above to **your own project** on github
2. Repeat the above steps for **your own project**.

Now that you have successfully run and debugged your own version of our project locally, you can push your code to github.com and then deploy your app to the internet.

- Please follow the instructions on Streamlit's website here:
    - https://docs.streamlit.io/streamlit-community-cloud/get-started/deploy-an-app

  In just a few short steps of pointing streamlit to **your own project's** GitHub URL main python app file, your app will be running live on the web!

- If you run into any problems like memory issues, as of the time of this writing, Streamlit might upgrade your machine for free If your app is for an educational or non-profit use case according to this.

Step 5: Optional: Deploying your own app to the web, hosted by Amazon AWS

- Same as the first part of Step 4: Fork the project and get your own project to run locally and then push to **your own project** on github.

- Deploying to AWS. There is no official way to do this yet. But all we really have to do is get Streamlit to run on a remote computer, just like how we got it to run our code on our local computer.
    - Streamlit is still developing its own docs on how to deploy to many different cloud services like AWS, Google GCP but they list some resources here:
        - https://discuss.streamlit.io/t/streamlit-deployment-guide-wiki/5099
    - We followed some of these more recent tutorials:
        - https://medium.com/@data.science.enthusiast/how-to-deploy-a-streamlit-machine-learning-app-over-aws-ec2-instance-12b6751268f1
        - https://www.linkedin.com/pulse/creating-shareable-data-apps-using-streamlit-aws-ec2-abhishek-gupta
    - The key steps we took are:
        - Sign up for an AWS account if you don't already have one.
            - **Important Note**: Although the account might be in the free tier, be mindful of costs and set up alerts in case you use something outside of the free tier.

- Set up an AWS EC2 instance (In our case we used Linux, 8GB of memory or more. **Again be sure to pay attention to costs!**).
- Log into the EC2 instance via Amazon's online terminal or use your own local terminal.
    - Launch a `tmux` session (so you can later exit the terminal and leave the server running) (Install tmux if not preinstalled on your chosen OS)
        - Type: `tmux new -s mysession`
    - Check out the code for your app so it is now on your EC2 instance.
        - Type: `git clone https://github/[YOUR_APP]`
    - Run the app:
        - Type: `streamlit run introduction.py`
        - Notice the IP address printed on the screen this time. It's not localhost like previously. This IP address is where your server is running.
        - Go to that IP address in your web browser. Your app is running on the web!
    - Exit the session by press CTRL-B and then press "d" key in your terminal
- Exit the terminal by typing: `exit`
- Your app is still running live on the web!

Congratulations! 🎉 You have successfully pulled the code, [potentially made code changes on your own], tested locally, and even pushed to git and deployed to the web live!

Jupyter Notebooks

This section talks about how to run our notebooks locally

Prerequisites

- You have cloned the project (explained in the **Search and recommender App** section above)

Step 0:

- You can check out and view our notebooks in GitHub. However not all of the data files are there.

- Instead you can view the same notebooks on DeepNote. Here is the link to our DeepNote notebooks: https://deepnote.com/workspace/test1-63d7-84aa8369-66c8-4270-ad20-b9f786b03094/project/Capstone-f2a50dc6-ff6a-45ff-9dbe-d7a35bd1e393/notebook/cluster_analysis-58dd11cd36a1402abbb5a265c87efc86
- You can download the entire notebooks and datasets there.

*** end of document ***