



# Đồ án tốt nghiệp Data Science

Topic: *Project Deployment*

[https://csc.edu.vn/data-science-machine-learning/do-an-tot-nghiep-data-science\\_310](https://csc.edu.vn/data-science-machine-learning/do-an-tot-nghiep-data-science_310)

2025



# Nội dung

---

1. Tổng quan Streamlit
2. Xây dựng ứng dụng Streamlit đầu tiên
3. Ứng dụng Data Science kết hợp Streamlit
4. Xuất bản ứng dụng lên Streamlit Cloud

# Tổng quan Streamlit

## □ Đặt vấn đề

- Bạn thực hiện dự án Data Science và khi hoàn thành dự án, bạn báo cáo kết quả đến các bộ phận liên quan. Bạn muốn người dùng cuối có thể vận dụng kết quả của bạn để dự báo/dự đoán một cách dễ dàng, không phải đọc code???

=> Bạn sẽ làm gì???

# Tổng quan Streamlit

## □ Giới thiệu Streamlit

- Streamlit là công cụ được xây dựng với mục đích dành cho Machine Learning Engineer, tạo ra giao diện web đơn giản, dễ sử dụng cho người dùng cuối => sản phẩm hoàn thiện, có tính ứng dụng cao.



<https://streamlit.io/>

# Tổng quan Streamlit

---

- Streamlit được biết đến như con đường nhanh nhất giúp đưa những bộ code của bạn thành một sản phẩm ứng dụng, “**The fastest way to build and share data apps**”
- Với bộ function sẵn có phong phú làm tăng khả năng tương tác trên app, Streamlit là công cụ lý tưởng cho những người muốn hướng tới xây dựng một sản phẩm hoàn thiện.

# Tổng quan Streamlit

## ❑ Đặc điểm của Streamlit

- Streamlit biến các data script thành các ứng dụng web có thể chia sẻ trong thời gian ngắn.
- Thư viện Python mã nguồn mở, miễn phí
- Xây dựng giao diện người dùng cho nhiều mục đích khác nhau, không giới hạn là ứng dụng dữ liệu hay máy học.
- Dễ thực hiện, không yêu cầu nhà phát triển có kinh nghiệm về giao diện

# Tổng quan Streamlit

## ❑ Vì sao nên chọn Streamlit?

- Không cần phải xử lý HTML/ CSS/ JavaScript
- Hỗ trợ định dạng văn bản Markdown text
- Có nhiều component/widget được tạo sẵn, giảm thời gian xây dựng giao diện
- Dễ dàng triển khai ứng dụng Streamlit lên Internet

# Nội dung

---

1. Tổng quan Streamlit
2. Xây dựng ứng dụng Streamlit đầu tiên
3. Ứng dụng Data Science kết hợp Streamlit
4. Xuất bản ứng dụng Streamlit Cloud

# Xây dựng ứng dụng Streamlit đầu tiên

## □ Cài đặt và sử dụng

- Visual Studio Code
- Streamlit
- Các thư viện khác liên quan đến Data Science/ Machine Learning nếu cần

*File hướng dẫn cài đặt dành cho cả 3 HĐH/ Colab được cung cấp*

## ❑ Các bước thực hiện

- Bước 1: Tạo folder làm việc (một project chỉ tạo 1 folder)
- Bước 2: Vào Visual Studio Code > File > Open Folder > chọn folder đã tạo ở Bước 1
- Bước 3: Trong folder được mở bên tay trái > New File > ten\_tap\_tin.py

# Xây dựng ứng dụng Streamlit đầu tiên

## ● Bước 4: Thực hiện việc viết code

- Import các thư viện cần thiết

```
import streamlit as st
```

...

- Viết code

- Để viết một dòng dạng title dùng:

```
st.title("Nội dung hiển thị")
```

- Để viết một dòng dạng header dùng:

```
st.header("Nội dung hiển thị")
```

# Xây dựng ứng dụng Streamlit đầu tiên

## ● Bước 5: Thực thi ứng dụng

- Click chuột phải > Chọn Run Code > hiển thị màn hình thực thi > chọn Terminal > gõ:

```
streamlit run ten_tap_tin.py
```

- Một trang web sẽ được mở và có thể xem kết quả, hoặc có thể click vào link được phát sinh sau khi thực thi thành công

# Xây dựng ứng dụng Streamlit đầu tiên

- Ví dụ: Tạo tập tin `demo_streamlit.py` có nội dung như sau rồi thực thi:

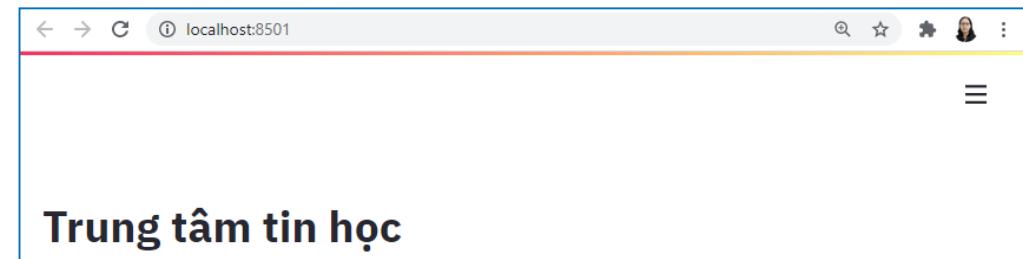
```
➊ demo_streamlit.py
 1 import streamlit as st
 2
 3 st.title("Trung tâm tin học")
```

**streamlit run demo\_streamlit.py**

You can now view your Streamlit app in your browser.

Local URL: <http://localhost:8501>

Network URL: <http://192.168.1.5:8501>



Bấm **Ctrl+C** để ngừng app

# Xây dựng ứng dụng Streamlit đầu tiên

localhost:8501

Menu

Capstone Project

## Trung Tâm Tin Học



Xe máy cũ

### Đồ án TN Data Science

Có 2 chủ đề trong khóa học:

- Topic 1: Dự đoán giá xe máy cũ, phát hiện xe máy bất thường
- Topic 2: Hệ thống gợi ý xe máy dựa trên nội dung, phân cụm xe máy

# Các Streamlit component cơ bản

- Streamlit hỗ trợ rất nhiều component, có thể tham khảo tại:

<https://docs.streamlit.io/develop/api-reference/widgets>,

<https://docs.streamlit.io/library/api-reference>

- Tạo multipage trong Streamlit app:

<https://docs.streamlit.io/get-started/tutorials/create-a-multipage-app>

# Nội dung

---

1. Tổng quan Streamlit
2. Xây dựng ứng dụng Streamlit đầu tiên
3. Ứng dụng Data Science kết hợp Streamlit
4. Xuất bản ứng dụng lên Streamlit Cloud

## □ Xây dựng ứng dụng

- Gợi ý xe máy trên Chợ Tốt

## □ Kết hợp Streamlit

- Xử lý, hiển thị kết quả, vẽ biểu đồ và hiển thị kết quả dự đoán

*Demo: Source code của project này đã được chia sẻ trên Google Classroom*

# Ứng dụng Data Science kết hợp Streamlit



Chọn xe máy bạn quan tâm:

Bán Vespa Sprint 125cc 2024 xanh dương, xe đẹp 95%

bike\_ID: [1](#)

Bạn vừa chọn:

**Bán Vespa Sprint 125cc 2024 xanh dương, xe đẹp 95%**

Thông tin:

Bán xe #Vespa Sprint 125cc. Mua mới tại #Topcom 01/2024, Xe chính chủ. Chạy được 14.000km. Giá: 66tr(TL). Anh/chị quan tâm nhắn Zalo số: \*\*\*. Công chứng Hợp đồng mua bán trong ngày. ...

Các xe máy khác bạn cũng có thể quan tâm:

Bán Xe Vespa Sprint gấp giá rẻ

bán cup125 xe như mới đầy đủ  
phụ kiện theo xe mới

CHÍNH CHỦ Cần bán Yamaha  
Exciter 2014 không côn

Description

Xe Vespa Piggio Sprint 2019  
mới 99% Xe nữ chạy còn rất  
mới. Minh không có nhu cầu  
đi xe bán lại giá tốt cho ai  
cần Xe biển số Sài Gòn, trùm  
mền để nhà chạy chỉ mới  
chạy 7.283km Ai cần ib bán

Description

bán cup125 xe như mới đầy  
đủ phụ kiện theo xe mới (   
lấy xe đen mới mua đưa qua  
) Do đổi xe đen nên nhượng  
lại xe xanh .chỉ để dành đi  
học . Cavet chính chủ bso

Description

Xe Yamaha Exciter không  
côn tay 4 số, máy êm, sang  
số mượt Dàn áo kiểu Exciter  
2014, màu xanh đen thể  
thao Giấy tờ đăng ký đầy đủ,  
biển số 92, chính chủ, sang

## ❑ Các bước thực hiện

- Tạo folder cho project
- Đưa các tập tin dữ liệu cần thiết vào folder
- Tạo file *content\_based\_app.py* và thực hiện phần xây dựng project kết hợp streamlit
- Tạo các tập tin deploy cần thiết
  - requirements.txt
  - setup.sh
  - Procfile

# Ứng dụng Data Science kết hợp Streamlit



- requirements.txt: là tập tin chứa các thư viện sử dụng cho Project, nằm trong thư mục gốc

The screenshot shows the Visual Studio Code interface. On the left is the Explorer sidebar with various project files listed. A red arrow points from the 'requirements.txt' file in the Explorer to the 'requirements.txt' file in the main editor area. The editor shows the following content:

```
1 streamlit
2 numpy
3 pandas
4 matplotlib
5 seaborn
6 scikit-learn
7 gensim
8 pyvi
```

# Ứng dụng Data Science kết hợp Streamlit



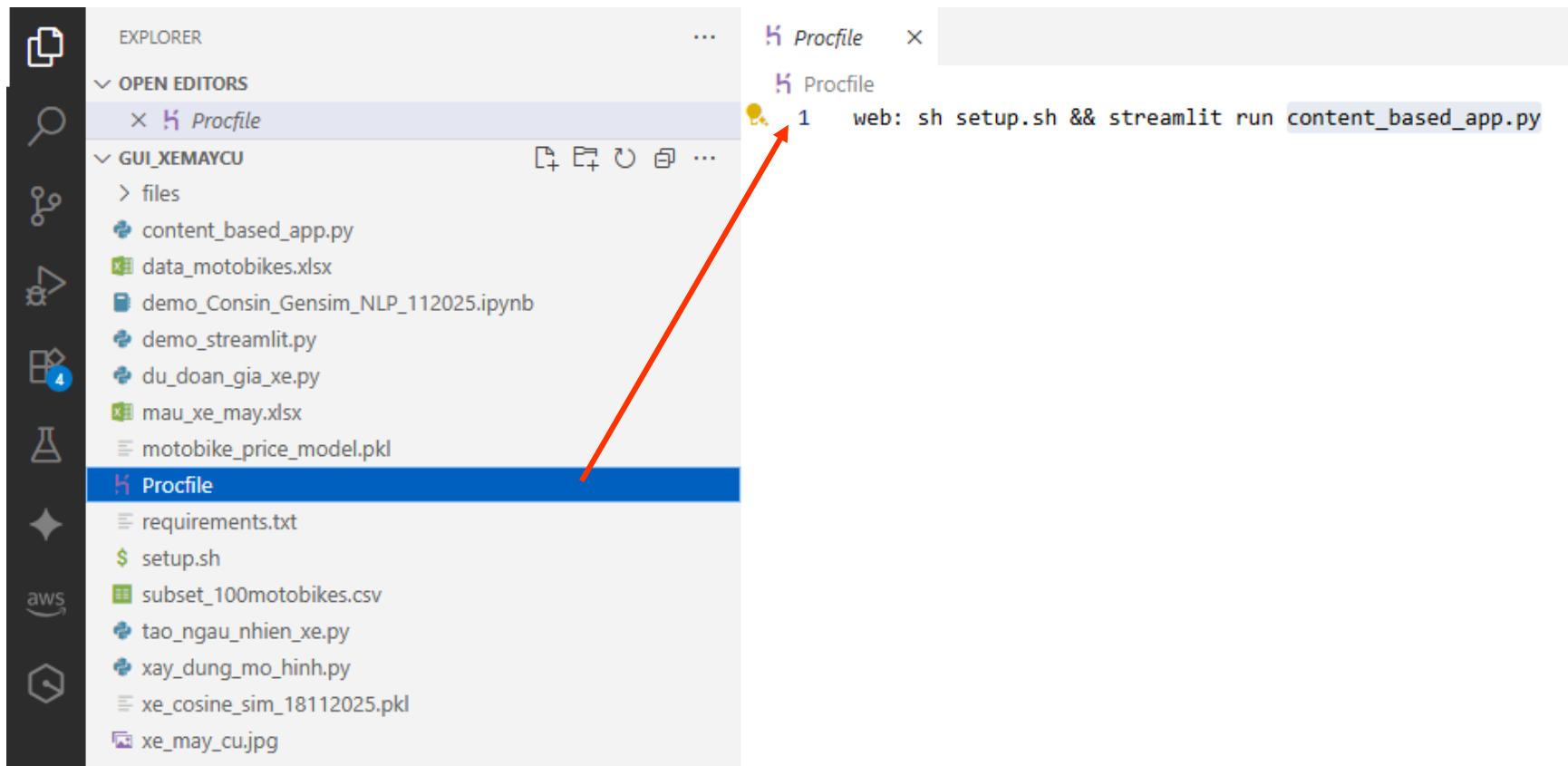
- setup.sh: nằm trong thư mục gốc

The screenshot shows the VS Code interface. On the left is the Explorer sidebar with a list of files and folders. A red arrow points from the 'setup.sh' entry in the Explorer to the code editor on the right. The code editor displays the contents of the 'setup.sh' script.

```
$ setup.sh
1 mkdir -p ~/.streamlit/
2 echo " \
3 [general]\n\
4 email = \"youremail@domain\"\n\
5 " > ~/.streamlit/credentials.toml
6 echo " \
7 [server]\n\
8 headless = true\n\
9 enableCORS=false\n\
10 port = $PORT\n\
11 " > ~/.streamlit/config.toml
12
```

# Ứng dụng Data Science kết hợp Streamlit

- Procfile (không có đuôi mở rộng): nằm trong thư mục gốc



# Nội dung

---

1. Tổng quan Streamlit
2. Xây dựng ứng dụng Streamlit đầu tiên
3. Ứng dụng Data Science kết hợp Streamlit
4. Xuất bản ứng dụng lên Streamlit Cloud

# Xuất bản ứng dụng lên Heroku

---

## ❑ Các bước thực hiện

- Thực hiện theo hướng dẫn nằm trong file [Xuat\\_ban\\_ung\\_dung\\_len\\_Streamlit\\_Cloud\\_tu\\_Github.PDF](#)

# Triển khai project theo Data Science Process



Các công việc cần thực hiện...

# Yêu cầu

- **Với 2 project đã triển khai, mỗi nhóm xây dựng cả 2 GUI (Project 1: Dự đoán giá xe, phát hiện bất thường; Project 2: Đề xuất xe máy dựa trên nội dung, phân cụm xe máy)**
- **Đưa ứng dụng GUI đã xây dựng lên Streamlit Cloud (hoặc host bất kỳ) => cung cấp link trang web vào file Excel đã share để được nhận xét (và chấm điểm)**
- **Chú ý: Với project 2 nếu kích thước quá lớn thì có thể giảm lượng dữ liệu khi build model trước khi làm GUI.**



# Yêu cầu

---

## ☐ Một GUI tối thiểu cần có:

- Menu (đặt bên trái trang hoặc ở trên):  
Business Problem, Evaluation & Report,  
New Prediction/ Analysis/  
Recommendation,...
- Nội dung tương ứng với mục được chọn  
trên menu.
- Thông tin nhóm thực hiện:
  - Họ tên HV 1, email, ...
  - ...

