CSCE 580: Introduction to AI Prof. Biplav Srivastava, Fall 2025

Finals / Dec4, 2025/ Instructions

- You have five days to complete the exam
- All the responses should be in your Github before ***the start of the class***. Submission will be in a folder called "**finals"**.
    - Use a readme.md in the folder if you like
    - You can have non-code responses uploaded as a .pdf. You can have same or different pdfs per question.
    - For code responses, created a subfolder (like ./finals/Q2/code) and upload it inside it.
    - We will use your readme.md to find the responses for the finals.
- For questions/ clarifications, ask Instructor biplav.s@sc.edu and TAs vishalp@email.sc.edu, kausik@email.sc.edu.

Total points = (20 + 80) = 100 points

Student Name:
Obtained =

DO EITHER of Q1a or Q1b

Q1a: [2 + 10 + 8 = 20 marks]

Context - We will test your understanding of graduate papers presented in class.
Instructions

1. The presentations by grad students are at: https://drive.google.com/drive/folders/1FMHzxO2JFwouZeeLjaN27EwiD-taPmU3?usp=sharing
2. Select one of the eight papers presented in the class by grad students. If you were a presenter, you must select a **paper different-one-than-yours**.

Question:

a) Write the name of the paper and student presenter you chose. **[2]**
b) Now, can you think and create a new example exemplifying the main conclusion of the paper. **[10]**
c) Describe how the conclusion is supported in your example. **[8]**

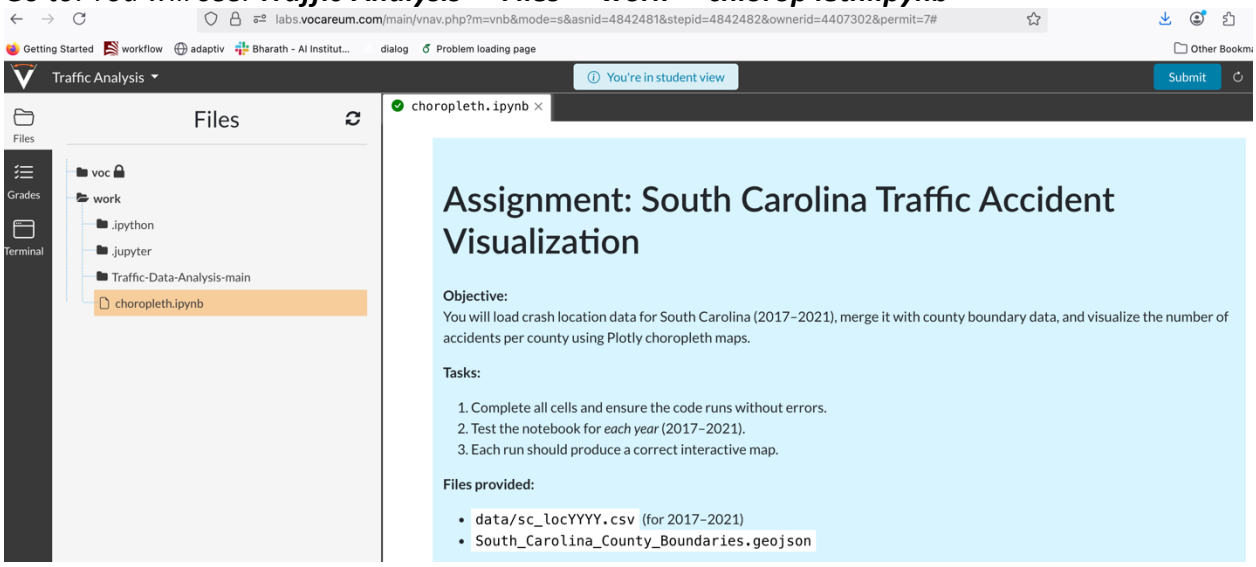
Q1b: [5 + 5 + 10 = 20 marks]

Context - We will have you experience real-world data science – including data, analysis and tools. Vocareum is a platform for online data science labs. We will ask you to run an online notebook using it and reproduce results. This is part of a NSF NAIRR funded resource for data science teaching we secured ***(****NAIRR250023: Privacy-Preserving Traffic Safety Analysis: A Module on Trusted AI With Real Data from Civic Partners****).***

Instructions:

Follow the detailed instructions in the attached sheet.
[*Summary*

1. *Go to: https://labs.vocareum.com/main/vnav.php?m=vnb&mode=t&asnid=4842481&stepid=4842482#*
2. *You will be asked to sign-in. Use single-signon using your USC's credentials*
3. *Go to: You will see:* ***Traffic-Analysis -> Files -> work -> chlorop leth.ipynb***



4. *Run the notbook till the end and see the resulting traffic analysis and visualization on SC map]*

<u>Questions / actions</u>:

a) If you were able to run the notebook till the end, list what data sources were used in the analysis. **[5]**

b) Put a copy of the final image for any year created in your response**.** **[5]**

c) We will verify Vocareum's backend for completion marks. **[10]**

Q2: Using AI for tackling a pressing teaching problem – classroom absenteeism. [80 points]

Context - We will test your understanding of concepts in our class setting. The students of CSCE 580 signed up their attendance on paper. We are making this data available as images. Your task is to come up with an image-based ***attendance audit system*** that takes these images and creates reports including number of classes held, median class attendance, dates with lowest and highest attendance, and correlation of attendance with course evaluations.

The class attendance data for 27 classes is available in image form at:
https://drive.google.com/drive/folders/1pKbZgxDDyc86Fmvkk1BGFljxuq5srFOp?usp=sharing

Instructions

1. Download the data and reduce size if needed.
2. Try to extract student information (name, username) and differentiate them from other data: e.g., class number, dates, and serial id. You can use pre-trained models (like LLMs – Llama/ Llava, Yolov5) or train your own. As a last resort, you can also manually tabulate and then analyze the data (although not recommended)
3. Think how you will check the correctness of the answers from your models.
4. Answer the questions asked about attendance.

Question:

a) Describe your data preparation, if any, and why or why not. **[20 points]**
b) Describe your steps to create a model – pre-trained, your own, manual **[30 points]**
c) Answer the questions from your analyses using the models
    a. What are the number of classes and their dates? **[10 points]**
    b. What is the median class attendance per class **[10 points]**
    c. What are the dates with lowest and highest attendance? **[10 points]**
    d. Is there a correlation of high attendance with course evaluations dates? When is the attendance highest **[10 points]**
d) If you had more time (say a week), what more could you have done to improve performance? **[10 points]**

Reference: The important course evaluation dates with in-room participation were

- Quiz 2: Oct 7
- Quiz 3: Nov 11
- Paper presentation (grad students only) : Nov 18