**BME 3053C Final Project**

**Proposal Due Date: Wednesday, November 13, at 11:59 PM**, on Canvas

**Spotlight Video Link Due Date: Tuesday, December 3, 11:59 PM**, on Canvas

**Final Report and Code Due Date: Wednesday, December 4, 11:59 PM**, on Canvas

**Introduction**

In this project, your goal is to propose a solution to a biomedical-related problem. Your team can choose a project topic of interest, such as EEG signal denoising, CT/MRI image restoration or registration, cell detection in microscopy images, outcome prediction using Electronic Health Records (EHR) data, or frequency analysis of 1D/2D signals. Your solution should involve training a machine learning model on an open-source dataset related to your topic. **You are expected to use AI assistants like ChatGPT and Cursor to assist you with the project.**

**Deliverables**

**Who needs to submit it?**

Just one group member is required to submit spotlight video link, project report, GitHub Link because it is set as a group assignment on Canvas. Every group member is required to complete the peer evaluation individually.

**What to submit?**

1. **Spotlight Video Link & PowerPoint Slides:** A YouTube link to your spotlight video of the presentation and code demo. Upload the PowerPoint slides used in the Spotlight video with the YouTube link on Canvas.
2. **Report:**Submit your report as aPDF File. Name the report file as **[TeamName]\_Report.pdf.**
3. **GitHub Repository:**
   * **Data:** Find an open-source dataset to train your chosen model. If the dataset is large, you can create a subset of the dataset (as long as you have enough data to train the model). Provide the dataset as a zip file in your repository or your Canvas submission. The data file should be named **[Team Name]\_Data.zip.**
   * **Code:** Please provide your code with appropriate comments that explain each step in your approach. **Your** **code should include a demo that others can run to replicate your approach.**
   * **ReadMe File:** A readme (.md) file with step-by-step instructions on how to run the code and where you found your dataset.
4. **Peer Evaluation Survey**: Please complete the peer evaluation survey on Canvas.

**Late/Penalty Policy**

You will be penalized if your code cannot run. The late submission policy applies universally, with no exception.

**Instructions**

**1. Spotlight Video:**

* Use Zoom to record a short video (~5 min) and upload to YouTube. Record it on your local computer, then upload it to YouTube. When uploading, you can set the video as “unlisted” so only people with the link can view it, and the video won’t be searchable.
* Include both your PowerPoint presentation and code demo in the Spotlight video.
* Use PowerPoint slides to introduce the project overview, problem, significance, data, methodology, and results.
  + PowerPoint slides should mainly contain figures and bulleted lists of text. Text font size should be at least 16.
* Walk users through how to use your code and show actual results on your data.
* The spotlight video can be presented by one team member or all members together (recommended).
* Include visual aids, videos, and demos where appropriate in your spotlight presentation, which will enhance your presentation.
* For the demo presentation, show the audience and the grader how to follow your steps to run your code.
* Showing stepwise intermediate results will make your presentation and demo more explainable and exciting.
* Upload the PowerPoint slides used in the spotlight video with the YouTube link on Canvas.
* See here for how to record a presentation in Zoom: <https://iu.pressbooks.pub/semesterchecklist/chapter/recording-an-individual-or-group-presentation-with-zoom/>

**2. Report:**

* Length: 4 pages, 11 Times New Rome, Single Spaced.
* Include all group member names in your final project report.
* **Problem Statement:** Clearly define the biomedical problem you are addressing.
* **Literature Review:** Discuss existing approaches and methodologies related to your problem. **You should have at least 10 citations in this section.** (Use the citation format of your choice)
* **Proposed Solution:** Describe your approach, including the machine learning model used, data preprocessing steps, and any feature engineering techniques applied.
* **Results and Evaluation:** Present your findings, including performance metrics, validation results, and any challenges encountered. Attach figures and/or tables to visualize your results.
* **Practical Application:** Explain how your solution can be applied in real-world biomedical settings and its potential impact.

**3. GitHub:**

* One member should create a GitHub repository and add the other members as collaborators. (Repository -> Settings -> Manage Access -> Invite a collaborator)
* If you choose to make your repository Public, you can just submit the link on Canvas.
* If you choose to make your repository Private, you can either turn it Public before submission (in Repository->Settings ->Options->Danger Zone->Change Repository Visibility) or add the instructor and STs as collaborators.
* Ensure that your code is well-documented and follows best practices for readability and reproducibility.

**Grading Rubric:**

|  |  |  |  |
| --- | --- | --- | --- |
| Category | Scoring Criteria | Total Points | Score |
| PowerPoint Slides  (10 pts) | The slides are simple yet informative | 10 |  |
| Video  (20 pts) | Project importance and impact | 5 |  |
| Method & Results | 5 |  |
| Presentation clarity and engagement | 10 |  |
| Report  (50 pts) | Significance of the research problem | 10 |  |
| Innovation of the research problem | 10 |  |
| A cited outline of related research and existing approaches | 10 |  |
| A sound and well-explained technical approach | 5 |  |
| Convincing results and well-motivated explanations | 10 |  |
| Logical summarization of the findings and overall writing | 5 |  |
| GitHub Repository  (20 pts) | The code runs well and produces the expected results | 10 |  |
| The code is intuitive to use and is thoroughly documented | 10 |  |
| Award | (Extra Credit) 1st, 2nd, and 3rd Place in Video Competition | 5 / 3 / 1 |  |
| Score | Total Points | 100 + 5 Extra |  |

**Potential Projects Ideas:**

1. **Cell Detection**

Accurate quantification of cells in biomedical images is essential for various research applications, including cancer diagnosis, drug development, and understanding cellular mechanisms. Traditional manual counting methods are labor-intensive and subject to significant inter-observer variability. In this project, you will develop an automated solution to detect and count cell nuclei in microscopy images using the Data Science Bowl 2018 dataset. This dataset provides well-annotated images from diverse cell types, enabling the development of robust models for cell nucleus detection and segmentation. By accurately segmenting nuclei, your solution will automate cell quantification, reduce observer variability, and enhance research efficiency across biomedical applications.

More Info: <https://www.kaggle.com/competitions/data-science-bowl-2018>

2. **Intracranial Hemorrhage Detection**

Intracranial hemorrhage, bleeding that occurs inside the cranium, is a serious health problem requiring rapid and often intensive medical treatment. For example, intracranial hemorrhages account for approximately 10% of strokes in the U.S., where stroke is the fifth-leading cause of death. Identifying the location and type of any hemorrhage present is a critical step in treating the patient.

Diagnosis requires an urgent procedure. When a patient shows acute neurological symptoms such as severe headache or loss of consciousness, highly trained specialists review medical images of the patient’s cranium to look for the presence, location, and type of hemorrhage. The process is complicated and often time consuming.

You’ll develop your solution using a rich image dataset provided by the Radiological Society of North America (RSNA®) in collaboration with members of the American Society of Neuroradiology and MD.ai.

More Info: <https://www.kaggle.com/c/rsna-intracranial-hemorrhage-detection/data>

**3. Optic Cup and Disc Ratio in Glaucoma Diagnosis**

Glaucoma is currently the leading reason of irreversible blindness in the world. It is commonly caused by elevated intraocular pressure (IOP), which causes mechanical straining and torsion of optic nerve and loss of retinal nerve fibers. Glaucoma changes the morphology of optic nerve head (ONH) and is usually characterized by the larger cup to disc ratio (CDR), pale disc, disc hemorrhage, etc. in the ONH region, while in OCT images thinning of retinal nerve fiber layer could be seen at the early stage.

The optic cup is the central cup-like area in the optic disc. The CDR is the comparison of the diameter of the cup to disc, which partially represents disease status. Determination of CDR varies among doctors and can be influenced by subjectivity. The normal CDR is 0.3 to 0.4. A larger CDR may indicate glaucoma or other diseases such as neuro-ophthalmic diseases. Previous studies showed that larger vertical (not horizontal) CDR is closely associated with progression of glaucoma. Although not accurate enough, CDR is very useful in clinical practice and evaluation of glaucoma.

Raw images and ground truth segmentation:

Challenge website: <https://idrid.grand-challenge.org/workshop/>

Data: <https://ieee-dataport.org/open-access/indian-diabetic-retinopathy-image-dataset-idrid>

More Info: <https://refuge.grand-challenge.org/Home/>

**4. Pneumonia Detection**

Pneumonia accounts for over 15% of all deaths of children under 5 years old internationally. In 2015, 920,000 children under the age of 5 died from the disease. In the United States, pneumonia accounts for over 500,000 visits to emergency departments and over 50,000 deaths in 2015, keeping the ailment on the list of top 10 causes of death in the country.

While common, accurately diagnosing pneumonia is a tall order. It requires review of a chest radiograph (CXR) by highly trained specialists and confirmation through clinical history, vital signs and laboratory exams. Pneumonia usually manifests as an area or areas of increased opacity on CXR. However, the diagnosis of pneumonia on CXR is complicated because of a number of other conditions in the lungs such as fluid overload (pulmonary edema), bleeding, volume loss (atelectasis or collapse), lung cancer, or post-radiation or surgical changes. Outside of the lungs, fluid in the pleural space (pleural effusion) also appears as increased opacity on CXR. When available, comparison of CXRs of the patient taken at different time points and correlation with clinical symptoms and history are helpful in making the diagnosis.

CXRs are the most commonly performed diagnostic imaging study. A number of factors such as positioning of the patient and depth of inspiration can alter the appearance of the CXR, complicating interpretation further. In addition, clinicians are faced with reading high volumes of images every shift.

Data and Info: <https://www.kaggle.com/c/rsna-pneumonia-detection-challenge>

**4. More potential projects**

Kaggle Health Competitions: <https://www.kaggle.com/tags/healthcare>

Kaggle Medicine Competitions: https://www.kaggle.com/tags/medicine

**Hint:** For your final project, you can use a subset of the dataset provided by the competitions online and modify the goal of the project to be practical and feasible based on what you learned from this course.

**5. Previous Student Projects**

* Detection of White Matter Atrophy Due to Beta-Amyloid Plaque Build-up in Alzheimer’s Disease
* Self-Identify Melanoma
* Glaucoma Diagnosis: Cup to Disk Ratio
* Analysis of FTIR Data for SLA 3D-Printed Polymers Used in Stapes Prosthetics
* Enumeration and Differentiation of White Blood Cells through Image Processing
* The Application of Statistical Parametric Mapping in Denoising and Enhancing BOLD Signal Contrast in fMRI Data
* Metaphase Spread Analysis
* Microscopy Image Analysis for Quantification of Cells via Segmentation
* Automotive Pneumonia Detection in X-ray Scans
* PPG Analysis as Accessible Cardiovascular Diagnosis
* 2D Cell Migration Analysis
* Detect Bone Fractures in X-Ray Images
* Analysis of Crosslinking in Polymers for Middle Ear Prosthetics