Final Project Proposal

Please make this document anonymous. Your team name should be anonymous.

Team name: ***KENN***

Note: when submitting this document to Gradescope, make sure to add all other team members to the submission. This can be done on the submission page after uploading (top right).

If you need to find team members, please use the thread under 'Final Project - Find Teammates' on Ed—pitch an idea!

Proposal Instructions

For your project proposal, please submit a one-to-two page document answering the questions below.

- What is your project idea?
  Our project is focused on developing a machine learning algorithm that can accurately distinguish between real and fake faces in digital images. By using a dataset of 10,000 real and fake faces, we aim to train and validate our algorithm to achieve high accuracy in detecting manipulated images. To enhance the diversity of the training data and improve the model's generalization ability, we will employ data augmentation techniques such as zooming, horizontal flipping, and brightness adjustment. Our model architecture includes convolutional layers for feature extraction and dense layers for classification.

- What is the socio-historical context that this project lives in?
  The socio-historical context of our project on real vs. fake face detection is closely tied to the recent rise of deepfake technology and image manipulation. Deepfake technology has been advancing rapidly in recent years, raising concerns about its potential misuse in spreading misinformation. It has been involved in malicious activities such as non-consensual deepfake pornography or impersonating public figures. As a result, there is a growing need for tools and techniques that can reliably distinguish between real and fake content to maintain trust in digital media. Our project aims to fulfill such needs.

- Please list three stakeholders that your project could impact, and describe how it could impact them.
  Social Media platforms: social media sites are riddled with fake images, however, there is no surefire way of identifying which posts are of real or fake people. The

classification of fake faces would provide a way to monitor fake faces so social media sites could warn users of or use this information to ban accounts.

Law enforcement: Deepfakes are used for identity theft and fraud, so with the capability to distinguish between face and real faces, violations of the law can be better enforced. This stakeholder can use the classification process to track people who are deceiving the general public of their identity, especially when it is harmful to others.

General public: People can easily be deceived by fake images/news, so with our project, we can give the public a tool to avoid such deception. As well, these deep fakes are particularly harmful when using another individual's likeness, which our project has the potential to identify.

- —–

- What are the skills of the team members? Conduct a skill assessment!
  Team Member 1: Python, Neural Networks, Machine Learning, Deep Learning
  Team Member 2: Python, Deep learning, machine learning
  Team Member 3: Python (+data scraping), Machine learning
  Team Member 4: Python, Machine learning

- What data will you use? For our data, we will use this dataset found on Kaggle: https://www.kaggle.com/datasets/sachchitkunichetty/rvf10k. This dataset consists of 7000 images, 3500 real faces, and 3500 fake faces.

- What software/hardware will you use?
  We plan to use existing machine/deep learning learning libraries, like tensorflow or pytorch, to implement our deep learning models. Since we are planning on implementing different models to see what approach works best, we will use the various types of models and layers (linear, max pooling, convolution, dense, etc) that these libraries provide in order to determine what comprises the best performing model is. The hardware we will use are just our personal computers to run these models.

- Who will do what? [For anonymity, please use ''"Team member 1 will..." or, alternatively, take on daring pseudonames.]
  Team Member 1: Data pre-processing and scraping (if necessary)
  Team Member 2: 1/2 layer model building
  Team Member 3: Optimization and possibly using other models. As well, experimenting with hyperparameters
  Team Member 4: Visualization and analysis of the results

- How will you know whether you have made progress? What will you measure?

1. Data collection (One or two datasets)
2. Data pre-processing (Can load and use data)
3. Model building
4. 1 layer, 2 layer, other types of models
5. Train the model
6. Optimization
7. Test the model
8. Analysis
9. Visualization of results, visualization of interpretation (remedy black box issue)

When testing the model, we will set a target accuracy and measure the accuracy.

- What technical problems do you foresee or have?

  Data

  - There can be minor problems regarding the way we load the data and labels (though unlikely)
  - It is also unclear if 10k examples from the linked dataset will be enough data for the model to learn from

  Building/Training/Testing model

  - Finding correct architecture for size and task of dataset may be challenging (since we are not reimplementing a paper)
  - During training there can be performance problems such as overfitting or underfitting. Training may not converge (e.g. loss gets stuck at local minimum)
  - Actual performance outside of dataset may be poor (e.g. in sample performance good, but on real life examples may not be able to detect well)

- Is there anything that we can do to help? E.G., resources, equipment.
  Direction for what types of deep and machine learning models are most effective for computer vision detection models would be helpful. Other equipment we may use are department machines for running these models.

Feel free to use these as paragraph headings, and also please include any media, references, etc.

After Proposal Submission

TA Assignment

After handing in your project proposal, your team will be assigned a TA to assist you. You should aim to meet with your TA once a week; this replaces TA office hours.

If you haven't heard from your TA a few days after the project proposal handin, please make a private Ed post and let us know which team you're on.

In your first meeting with your TA, your goal is to have your idea sanity-checked:

- Do you actually have the data?
- Do you actually have the compute?
- Is there code you need but don't have access to?
- Is there an area where you need help?

Some of these things will be outlined in your proposals, but talking through it with your TA as soon as possible will help you find potential road blocks and get the ball rolling.