[9/7/2021] MQP Meeting Minutes

Attendees:

* Ryan Astor (Note Taker)
* Kyle Costello
* Alek Lewis
* Josh DeOliviera
* Walter Gerych
* Elke Rundensteiner

Minutes:

* Listed accomplishments
* Does the team know each other?
  + Strengths weaknesses?
  + The team has worked well together so far
* Understand PyTorch
  + Fairly intuitive
  + Will have to look back for syntax
  + Has to read a few times for a full comprehension
* Original GAN paper
  + What does the ABCD figure mean?
    - Original figure from the paper
    - Shows how the input data is mapped to the distribution of data
    - The blue line shows the guess of getting it right from the discriminator
    - Blackline is the original data
    - Greenline is generated data from random noise
    - Y shows the distribution
  + Strengths and weaknesses of GANs
  + What is meant by function
    - Looking at all of the loss functions
  + Does anything need to be done with different loss
    - GANs are trying to fool a discriminator
    - Many kinds can be used
    - Leaky Relu with some values, not 0 in negative
    - Binary cross-entropy is possible to be used too
* Conditional GANs
  + GANs converge to the dominant class because it is easier to do so
  + To make all of the classes we use conditional GANs to specify the wanted output
  + Discriminator now has 2 jobs, discriminate real vs fake data and make sure it is the correct class
  + Probability is between 1 and 0 where 1 is it is 100% sure that it is both the right class and real data and 0 is ambiguity on if the class is wrong or it is fake of it is both
  + **Make the slide 6 figure a bit more clear by changing the colors up a bit or some other way to fix it**
  + **2014 date is also wrong on slide 6, double-check the data**
* Controllable GAN
  + Is this about unbalanced data?
    - Good for classifications of a more specific type than the previous paper
  + What about super unbalanced like 99.99% dominant class?
    - The real data is more like 70% 30% so this is more realistic end easier to classify
  + Generation should always be of real data
  + Low confidence creates ambiguity in the old GAN
    - This new one separates the concerns into 2 probabilities so they can be asked and viewed separately
* Explicitly Controllable GANs
  + Takes vectors for each desired feature and encodes them into something that the generator can use
  + Can make data in a completely arbitrary matter
* Skipped many slides
* Next steps
  + Is there a good way to test each other on training
    - We can talk about it as a group
  + What does the novel training mean
    - We can start to look for them
    - Could be a bit hidden
* Takeaways
  + **Add Prof Rundensteiner to the GitHub**
  + **Make it private**